

VLM Multi-Model Comparison Report

Dataset: radiation_treatment_outcomes_v1_1

Subtask: radiation_outcome

Question:

Based on the provided medical history and cancer evidence, what is the patient's expected response to external beam radiation therapy?

Correct Answer: Insufficient follow-up or missing data

Model: OpenGVLab/InternVL3_5-8B

all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient is diagnosed with stage IIE nodular sclerosing Hodgkin's disease involving the right hilum and left supraclavicular region. The timeline indicates that she is undergoing a course of Stanford 5 chemotherapy followed by consolidative involved-field radiotherapy. The plan includes a treatment simulation and PET CT scan to assess the extent of disease before starting radiotherapy. The timeline does not provide information on the outcome of the radiotherapy, such as whether the tumor symptoms were controlled, returned, or if a complete response was achieved. Therefore, based on the provided information, the expected response to external beam radiation therapy cannot be definitively determined from the timeline alone. However, the absence of documented evidence of tumor response post-treatment suggests that the label "Never achieved complete response" might be considered, but this is speculative without explicit evidence of treatment outcome. | Never achieved complete response

axial_all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient is diagnosed with stage IIE nodular sclerosing Hodgkin's disease involving the right hilum and left supraclavicular region. The timeline indicates that she is undergoing a course of Stanford 5 chemotherapy followed by consolidative involved-field radiotherapy. The plan includes a treatment simulation and PET CT scan to assess the extent of disease before starting radiotherapy. The timeline does not provide information on the outcome of the radiotherapy, such as whether the tumor symptoms were controlled, returned, or if a complete response was achieved. Therefore, based on the provided information, the expected response to external beam radiation therapy cannot be definitively determined from the timeline alone. However, the absence of documented evidence of tumor response post-treatment suggests that the label "Never achieved complete response" might be considered, but this is speculative without explicit evidence of treatment outcome. | Never achieved complete response

no_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient is diagnosed with stage IIE nodular sclerosing Hodgkin's disease involving the right hilum and left supraclavicular region. The timeline indicates that she is undergoing a course of Stanford 5 chemotherapy followed by consolidative involved-field radiotherapy. The plan includes a treatment simulation and PET CT scan to assess the extent of disease before

starting radiotherapy. The timeline does not provide information on the outcome of the radiotherapy, such as whether the tumor symptoms were controlled, returned, or if a complete response was achieved. Therefore, based on the provided information, the expected response to external beam radiation therapy cannot be definitively determined from the timeline alone. However, the absence of documented evidence of tumor response post-treatment suggests that the label "Never achieved complete response" might be considered, but this is speculative without explicit evidence of treatment outcome. | Never achieved complete response

sagittal_all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label> The patient is diagnosed with stage IIE nodular sclerosing Hodgkin's disease involving the right hilum and left supraclavicular region. The timeline indicates that she is undergoing a course of Stanford 5 chemotherapy followed by consolidative involved-field radiotherapy. The plan includes a treatment simulation and PET CT scan to assess the extent of disease before starting radiotherapy. The timeline does not provide information on the outcome of the radiotherapy, such as whether the tumor symptoms were controlled, returned, or if a complete response was achieved. Therefore, based on the provided information, the expected response to external beam radiation therapy cannot be definitively determined from the timeline alone. However, the absence of documented evidence of tumor response post-treatment suggests that the label "Never achieved complete response" might be considered, but this is speculative without explicit evidence of treatment outcome. | Never achieved complete response

Model: OctoMed/OctoMed-7B

all_image: [INCORRECT] <think>First, I carefully examine the entire patient timeline to identify key points relevant to the question about the expected response to external beam radiation therapy. The patient has Hodgkin's disease, specifically nodular sclerosis subtype, with stage IIE disease involving the right hilum and left supraclavicular region. The timeline indicates that the patient underwent chemotherapy (Stanford 5) and then was scheduled for consolidative involved-field radiotherapy after completing chemotherapy. The PET CT performed before starting radiotherapy showed active disease in those areas. The patient's physical exam and lab results show stable vital signs and no acute distress, but there are ongoing side effects from chemotherapy such as rash, nausea, etc. Importantly, the patient's treatment plan includes radiotherapy targeting the bulky sites identified on PET CT. However, the timeline does not provide any information indicating whether the radiotherapy was completed or its outcome. There are no follow-up reports or documentation of response assessment post-radiotherapy. The last entries are from June 2010, well before any potential completion date of radiotherapy, suggesting that the radiotherapy itself may not have been administered yet or the outcome is unknown. Given the lack of data on the actual administration or outcome of the radiation therapy, it is impossible to determine if the tumor symptoms were controlled, returned, or if a complete response was achieved. Therefore, based solely on the provided timeline, the most accurate conclusion is that the patient's expected response cannot be determined because the radiotherapy treatment hasn't been completed or evaluated yet.</think> <brief reasoning trace with explanation for label selection> | <label> </label></think> \boxed{Never achieved complete response}

axial_all_image: [INCORRECT] <think>First, I carefully examine the entire patient timeline to identify key points relevant to the question about the expected response to external beam radiation therapy. The patient has Hodgkin's disease, specifically nodular sclerosis subtype, with stage IIE disease involving the right hilum and left supraclavicular region. The timeline indicates that the patient underwent chemotherapy (Stanford 5) and then was scheduled for consolidative involved-field radiotherapy after completing chemotherapy. The PET CT performed before starting radiotherapy showed active disease in those areas. The patient's physical exam and lab results show stable vital signs and no acute distress, but there are ongoing side effects from chemotherapy such as rash, nausea, etc. Importantly, the patient's

treatment plan includes radiotherapy targeting the bulky sites identified on PET CT. However, the timeline does not provide any information indicating whether the radiotherapy was completed or its outcome. There are no follow-up reports or documentation of response assessment post-radiotherapy. The last entries are from June 2010, well before any potential completion date of radiotherapy, suggesting that the radiotherapy itself may not have been administered yet or the outcome is unknown. Given the lack of data on the actual administration or outcome of the radiation therapy, it is impossible to determine if the tumor symptoms were controlled, returned, or if a complete response was achieved. Therefore, based solely on the provided timeline, the most accurate conclusion is that the patient's expected response cannot be determined because the radiotherapy treatment hasn't been completed or evaluated yet.</think> <brief reasoning trace with explanation for label selection> | <label></label></think> \boxed{Never achieved complete response}

no_image: [INCORRECT] <think>First, I carefully examine the entire patient timeline to identify key points relevant to the question about the expected response to external beam radiation therapy. The patient has Hodgkin's disease, specifically nodular sclerosis subtype, with stage IIIE disease involving the right hilum and left supraclavicular region. The timeline indicates that the patient underwent chemotherapy (Stanford 5) and then was scheduled for consolidative involved-field radiotherapy after completing chemotherapy. The PET CT performed before starting radiotherapy showed active disease in those areas. The patient's physical exams and lab results show stable vital signs and no acute distress, but there are ongoing side effects from chemotherapy such as rash, nausea, etc. Importantly, the patient's treatment plan includes radiotherapy targeting the bulky sites identified earlier. However, the timeline does not provide any information indicating whether the radiotherapy was completed or its outcome. There are no follow-up reports or documentation of response assessment post-radiotherapy. The last entries are from June 2010, well before any potential completion date of radiotherapy, suggesting that the radiotherapy itself may not have been administered yet or the outcome is unknown. Given that the question asks about the expected response to external beam radiation therapy based solely on the provided timeline, and considering that the timeline lacks data on post-radiotherapy evaluation (such as restaging scans showing response), it is impossible to determine if the tumor symptoms were controlled, returned, or if a complete response was achieved. Therefore, the only appropriate conclusion is that the patient never achieved a complete response because the radiotherapy outcome is not documented here.</think> <brief reasoning trace with explanation for label selection> | Never achieved complete response The patient's timeline documents ongoing treatment with chemotherapy followed by planned consolidative radiotherapy. However, there is no evidence in the provided records of post-radiotherapy assessment confirming symptom control, recurrence, or complete response. The absence of follow-up data or documented outcomes necessitates selecting "Never achieved complete response" as the most accurate label based on the available information.

<answer>Never achieved complete response</answer>

sagittal_all_image: [INCORRECT] <think>First, I carefully examine the entire patient timeline to identify key points relevant to the question about the expected response to external beam radiation therapy. The patient has Hodgkin's disease, specifically nodular sclerosis subtype, with stage IIIE disease involving the right hilum and left supraclavicular region. The timeline indicates that the patient underwent chemotherapy (Stanford 5) and then was scheduled for consolidative involved-field radiotherapy after completing chemotherapy. The PET CT performed before starting radiotherapy showed active disease in those areas. The patient's physical exam and lab results show stable vital signs and no acute distress, but there are ongoing side effects from chemotherapy such as rash, nausea, etc. Importantly, the patient's treatment plan includes radiotherapy targeting the bulky sites identified on PET CT. However, the timeline does not provide any information indicating whether the radiotherapy was completed or its outcome. There are no follow-up reports or documentation of response assessment post-radiotherapy. The last entries are from June 2010, well before any potential

completion date of radiotherapy, suggesting that the radiotherapy itself may not have been administered yet or the outcome is unknown. Given the lack of data on the actual administration or outcome of the radiation therapy, it is impossible to determine if the tumor symptoms were controlled, returned, or if a complete response was achieved. Therefore, based solely on the provided timeline, the most accurate conclusion is that the patient's expected response cannot be determined because the radiotherapy treatment hasn't been completed or evaluated yet.</think> <brief reasoning trace with explanation for label selection> | <label></label></think> \boxed{Never achieved complete response}

Model: Qwen/Qwen3-VL-8B-Instruct

all_image: [INCORRECT] The patient's timeline indicates she was diagnosed with stage IIE BX nodular sclerosing Hodgkin's lymphoma involving the right hilum and supraclavicular region, and was planned to receive consolidative involved-field radiotherapy following chemotherapy. The timeline documents that she was scheduled to begin radiation therapy 1–3 weeks after completing chemotherapy (last dose on 06/17/2010), and treatment planning (including PET/CT and simulation) was initiated in June 2010. However, the timeline contains no evidence of post-treatment imaging, clinical follow-up, or documentation of tumor response (e.g., symptom control, recurrence, or complete response) after radiation therapy. The only documented outcomes are pre-treatment staging and chemotherapy side effects, with no post-radiation assessment. Therefore, based on the provided evidence, the expected response to radiation therapy cannot be determined as "controlled" or "returned," and there is no indication of a complete response. The correct label is "Never achieved complete response" because no post-treatment evidence of response is present. <brief reasoning trace with explanation for label selection> | Never achieved complete response

axial_all_image: [INCORRECT] The patient's timeline indicates she was diagnosed with stage IIE BX nodular sclerosing Hodgkin's lymphoma involving the right hilum and supraclavicular region, and was planned to receive consolidative involved-field radiotherapy following chemotherapy. The timeline documents that she was scheduled to begin radiation therapy 1–3 weeks after completing chemotherapy (last dose on 06/17/2010), and treatment planning (including PET/CT and simulation) was initiated in June 2010. However, the timeline contains no evidence of post-treatment imaging, clinical follow-up, or documentation of tumor response (e.g., symptom control, recurrence, or complete response) after radiation therapy. The only documented outcomes are pre-treatment staging and chemotherapy side effects, with no post-radiation assessment. Therefore, based on the provided evidence, the expected response to radiation therapy cannot be determined as "controlled" or "returned," and there is no indication of a complete response. The correct label is "Never achieved complete response" because no post-treatment evidence of response is present. <brief reasoning trace with explanation for label selection> | Never achieved complete response

no_image: [INCORRECT] The patient's timeline indicates she was diagnosed with stage IIE BX nodular sclerosing Hodgkin's lymphoma involving the right hilum and supraclavicular region, and was planned to receive consolidative involved-field radiotherapy following chemotherapy. The timeline documents that she was scheduled to begin radiation therapy 1–3 weeks after completing chemotherapy (last chemo dose: 06/17/2010), and treatment planning (including PET/CT and simulation) was initiated in June 2010. However, the timeline contains no documentation of post-treatment imaging, clinical follow-up, or evidence of tumor response or recurrence. There is no mention of tumor symptoms returning or achieving complete response. The only documented clinical status is that she denied symptoms (cough, night sweats, weight loss, etc.) at the time of chemotherapy completion, but no post-radiation outcome is recorded. Given the absence of any post-radiation outcome data in the timeline, and the fact that the timeline ends with treatment delivery (CPT 77413, 77414, 77427) without follow-up, we cannot confirm tumor control or recurrence. Therefore, the only valid inference is that the timeline does

not provide evidence of tumor response or recurrence — it does not state “Tumor symptoms controlled” or “Tumor symptoms returned,” and “Never achieved complete response” is the only option that aligns with the lack of documented evidence of response. <brief reasoning trace with explanation for label selection> | Never achieved complete response

sagittal_all_image: [INCORRECT] The patient's timeline indicates she was diagnosed with stage IIE BX nodular sclerosing Hodgkin's lymphoma involving the right hilum and supraclavicular region, and was planned to receive consolidative involved-field radiotherapy following chemotherapy. The timeline documents that she was scheduled to begin radiation therapy 1–3 weeks after completing chemotherapy (last dose on 06/17/2010), and treatment planning (including PET/CT and simulation) was initiated in June 2010. However, the timeline contains no evidence of post-treatment imaging, clinical follow-up, or documentation of tumor response (e.g., symptom control, recurrence, or complete response) after radiation therapy. The only documented outcomes are pre-treatment staging and chemotherapy side effects, with no post-radiation assessment. Therefore, based on the provided evidence, the expected response to radiation therapy cannot be determined as “controlled” or “returned,” and there is no indication of a complete response. The correct label is “Never achieved complete response” because no post-treatment evidence of response is present. <brief reasoning trace with explanation for label selection> | Never achieved complete response

Model: google/medgemma-1.5-4b-it

all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>The patient has a history of nodular sclerosing Hodgkin's lymphoma involving the right hilum and supraclavicular region. The initial diagnosis was stage IIE BX. The patient completed a course of chemotherapy (Stanford 5) and is due to start consolidative radiotherapy. The PET CT scan performed on 02/25/2010 showed a large mass in the right hilum (9 x 8 x 6 cm) with an SUV of 12 and a hypermetabolic left supraclavicular lymph node (1.8 cm). The pathology report from the biopsy on 03/11/2010 confirmed nodular sclerosing Hodgkin's lymphoma. The patient is currently at week 9 of her 12-week chemotherapy course. The plan is to start consolidative radiotherapy to the bulky sites (right hilum and right parenchyma) 1 to 3 weeks after completing chemotherapy. The question asks about the expected response to external beam radiation therapy. Given the initial stage (IIE BX) and the presence of bulky disease identified on the PET CT scan, the patient is expected to have a limited response to radiation therapy. The disease is already present in multiple sites, including bulky areas, and the patient has already undergone chemotherapy. Therefore, it is unlikely that radiation therapy alone will achieve a complete response. <label>Never achieved complete response</label>

axial_all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>The patient has a history of nodular sclerosing Hodgkin's lymphoma involving the right hilum and supraclavicular region. The initial diagnosis was stage IIE BX. The patient completed a course of chemotherapy (Stanford 5) and is due to start consolidative radiotherapy. The PET CT scan performed on 02/25/2010 showed a large mass in the right hilum (9 x 8 x 6 cm) with an SUV of 12 and a hypermetabolic left supraclavicular lymph node (1.8 cm). The pathology report from the biopsy on 03/11/2010 confirmed nodular sclerosing Hodgkin's lymphoma. The patient is currently at week 9 of her 12-week chemotherapy course. The plan is to start consolidative radiotherapy to the bulky sites (right hilum and right parenchyma) 1 to 3 weeks after completing chemotherapy. The question asks about the expected response to external beam radiation therapy. Given the initial stage (IIE BX) and the presence of bulky disease identified on the PET CT scan, the patient is expected to have a limited response to radiation therapy. The disease is already present in multiple sites, including bulky areas, and the patient has already undergone chemotherapy. Therefore, it is unlikely that radiation therapy alone will achieve a complete response. <label>Never achieved complete response</label>

no_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient has a history of nodular sclerosing Hodgkin's lymphoma involving the right hilum and supraclavicular region. The initial diagnosis was stage IIE BX. The patient completed a course of chemotherapy (Stanford 5) and is due to start consolidative radiotherapy. The PET CT scan performed on 02/25/2010 showed a large mass in the right hilum (9 x 8 x 6 cm) with an SUV of 12 and a hypermetabolic left supraclavicular lymph node (1.8 cm). The pathology report from the biopsy on 03/11/2010 confirmed nodular sclerosing Hodgkin's lymphoma. The patient is currently at week 9 of her 12-week chemotherapy course. The plan is to start consolidative radiotherapy to the bulky sites (right hilum and right parenchyma) 1 to 3 weeks after completing chemotherapy. The question asks about the expected response to external beam radiation therapy. Given the initial stage (IIE BX) and the presence of bulky disease identified on the PET CT scan, the patient is expected to have a limited response to radiation therapy. The disease is already present in multiple sites, including bulky areas, and the patient has already undergone chemotherapy. Therefore, it is unlikely that radiation therapy alone will achieve a complete response. <label>Never achieved complete response</label>

sagittal_all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient has a history of nodular sclerosing Hodgkin's lymphoma involving the right hilum and supraclavicular region. The initial diagnosis was stage IIE BX. The patient completed a course of chemotherapy (Stanford 5) and is due to start consolidative radiotherapy. The PET CT scan performed on 02/25/2010 showed a large mass in the right hilum (9 x 8 x 6 cm) with an SUV of 12 and a hypermetabolic left supraclavicular lymph node (1.8 cm). The pathology report from the biopsy on 03/11/2010 confirmed nodular sclerosing Hodgkin's lymphoma. The patient is currently at week 9 of her 12-week chemotherapy course. The plan is to start consolidative radiotherapy to the bulky sites (right hilum and right parenchyma) 1 to 3 weeks after completing chemotherapy. The question asks about the expected response to external beam radiation therapy. Given the initial stage (IIE BX) and the presence of bulky disease identified on the PET CT scan, the patient is expected to have a limited response to radiation therapy. The disease is already present in multiple sites, including bulky areas, and the patient has already undergone chemotherapy. Therefore, it is unlikely that radiation therapy alone will achieve a complete response. <label>Never achieved complete response</label>

Model: google/medgemma-4b-it

all_image: [INCORRECT] Tumor symptoms controlled The patient has completed 12 weeks of Stanford 5 chemotherapy and is due to start consolidative radiotherapy. The PET CT from February 2010 shows a large mass in the right hilum, extending into the hilar region, associated with severe encasement of the right lower lobe and right middle lobe bronchi, and right middle lobe. There is a 1.8 cm hypermetabolic left supraclavicular lymph node. The patient's symptoms are controlled, and she is due to start radiotherapy. choice

axial_all_image: [INCORRECT] Tumor symptoms controlled The patient has completed 12 weeks of Stanford 5 chemotherapy and is due to start consolidative radiotherapy. The PET CT from February 2010 shows a large mass in the right hilum, extending into the hilar region, associated with severe encasement of the right lower lobe and right middle lobe bronchi, and right middle lobe. There is a 1.8 cm hypermetabolic left supraclavicular lymph node. The patient's symptoms are controlled, and she is due to start radiotherapy. choice

no_image: [INCORRECT] Tumor symptoms controlled The patient has completed 12 weeks of Stanford 5 chemotherapy and is due to start consolidative radiotherapy. The PET CT from February 2010 shows a large mass in the right hilum, extending into the hilar region, associated with severe encasement of the right lower lobe and right middle lobe bronchi, and right middle lobe. There is a 1.8 cm hypermetabolic left supraclavicular lymph node. The patient's symptoms are controlled, and she is due to start radiotherapy. choice

sagittal_all_image: [INCORRECT] Tumor symptoms controlled The patient has completed 12 weeks of Stanford 5 chemotherapy and is due to start consolidative radiotherapy. The PET CT from February 2010 shows a large mass in the right hilum, extending into the hilar region, associated with severe encasement of the right lower lobe and right middle lobe bronchi, and right middle lobe. There is a 1.8 cm hypermetabolic left supraclavicular lymph node. The patient's symptoms are controlled, and she is due to start radiotherapy. choice

Dataset: tb_v1_1_tb_classification_tasks

Subtask: pneumonitis_infection_answer

Question:

Based on the provided medical history, cancer evidence, and imaging findings, which diagnosis is most consistent with the imaging differential?

Correct Answer: Infection

Model: OpenGVLab/InternVL3_5-8B

all_image: [INCORRECT] <The patient's medical history includes a diagnosis of malignant neoplasm of the trachea and unspecified bronchus and lung, indicating a primary lung cancer. The imaging findings, specifically the computed tomography (CT) scans of the thorax and abdomen/pelvis, are likely to show evidence of cancer progression. The presence of multiple imaging studies and the specific diagnosis of cancer suggest that the imaging findings are consistent with cancer progression rather than pneumonitis, infection, or neither.> | <Cancer Progression>

axial_all_image: [INCORRECT] <The patient's medical history includes a diagnosis of malignant neoplasm of the trachea and unspecified bronchus and lung, indicating a primary lung cancer. The imaging findings, specifically the computed tomography (CT) scans of the thorax and abdomen/pelvis, are likely to show evidence of cancer progression. The presence of multiple imaging studies and the specific diagnosis of cancer suggest that the imaging findings are consistent with cancer progression rather than pneumonitis, infection, or neither.> | <Cancer Progression>

no_image: [INCORRECT] <The patient's medical history includes a diagnosis of malignant neoplasm of the trachea and unspecified bronchus and lung, indicating a primary lung cancer. The imaging findings, specifically the computed tomography (CT) scans of the thorax and abdomen/pelvis, are likely to show evidence of cancer progression. The presence of multiple imaging studies and the specific diagnosis of cancer suggest that the imaging findings are consistent with cancer progression rather than pneumonitis, infection, or neither.> | <Cancer Progression>

sagittal_all_image: [INCORRECT] <The patient's medical history includes a diagnosis of malignant neoplasm of the trachea and unspecified bronchus and lung, indicating a primary lung cancer. The imaging findings, specifically the computed tomography (CT) scans of the thorax and abdomen/pelvis, are likely to show evidence of cancer progression. The presence of multiple imaging studies and the specific diagnosis of cancer suggest that the imaging findings are consistent with cancer progression rather than pneumonitis, infection, or neither.> | <Cancer Progression>

Model: OctoMed/OctoMed-7B

all_image: [INCORRECT] <think>First, I need to analyze the patient timeline carefully to understand their medical history and current conditions. The patient has several diagnoses listed: - Hypothyroidism (E03.9) – managed with levothyroxine. - Abdominal aortic aneurysm (I71.4) – non-ruptured. - Malignant neoplasm of overlapping sites of unspecified bronchus and

lung (C34.80) – this suggests possible lung or airway malignancy. - Malignant neoplasm of trachea (C33) – indicates another primary or secondary cancer site involving the trachea. - Other nonspecific abnormal lung findings (R91.8) – could relate to infection, inflammation, or malignancy. - Imaging studies: CT scans of thorax and abdomen/pelvis with contrast were performed. - Case management visit noted. The question asks about the imaging differential diagnosis based on these data. The options include pneumonitis, infection, both, cancer progression, or neither. Key points from the timeline: - The presence of two malignant neoplasms (lung/bronchus and trachea) strongly suggests advanced or multifocal cancer. - The CT scans (thorax and abdomen/pelvis) would have been used to evaluate these malignancies. - The nonspecific lung abnormality (R91.8) might indicate an inflammatory process like pneumonitis or infection, but it's not definitive. - There is no mention of biopsy results, lab markers (like white blood cell count), or symptoms such as fever or cough that would support infection. - Pneumonitis can occur due to radiation therapy or chemotherapy, but there's no indication here of recent treatments mentioned. - Cancer progression is plausible given the known malignancies and the imaging findings likely showing tumor burden or spread. Since the timeline does not provide direct evidence of infection (e.g., positive cultures, elevated inflammatory markers) or pneumonitis (e.g., post-radiation changes confirmed), but does show multiple cancers, the most consistent diagnosis with the imaging differential is cancer progression.</think> <brief reasoning trace with explanation for label selection> | <label> The patient has documented malignant neoplasms of the lung/bronchus and trachea (C34.80 and C33), supported by CT imaging (CPT4 codes). The nonspecific lung abnormality (R91.8) lacks specific evidence pointing to infection or pneumonitis (no lab results, symptoms, or treatment details). Given the known malignancies and imaging findings, cancer progression is the most consistent diagnosis. <answer>Cancer Progression</answer></think>

axial_all_image: [INCORRECT] <think>First, I need to analyze the patient timeline carefully to understand their medical history and current conditions. The patient has several diagnoses listed: - Hypothyroidism (E03.9) – managed with levothyroxine. - Abdominal aortic aneurysm (I71.4) – non-ruptured. - Malignant neoplasm of overlapping sites of unspecified bronchus and lung (C34.80) – this suggests possible lung or airway malignancy. - Malignant neoplasm of trachea (C33) – indicates another primary or secondary cancer site involving the trachea. - Other nonspecific abnormal lung findings (R91.8) – could relate to infection, inflammation, or malignancy. - Imaging studies: CT scans of thorax and abdomen/pelvis with contrast were performed. - Case management visit noted. The question asks about the imaging differential diagnosis based on these data. The options include pneumonitis, infection, both, cancer progression, or neither. Key points from the timeline: - The presence of two malignant neoplasms (lung/bronchus and trachea) strongly suggests advanced or multifocal cancer. - The CT scans (thorax and abdomen/pelvis) would have been used to evaluate these malignancies. - The nonspecific lung abnormality (R91.8) might indicate an inflammatory process like pneumonitis or infection, but it's not definitive. - There is no mention of biopsy results, lab markers (like white blood cell count), or symptoms such as fever or cough that would support infection. - Pneumonitis can occur due to radiation therapy or chemotherapy, but there's no indication here of recent treatments mentioned. - Cancer progression is plausible given the known malignancies and the imaging findings likely showing tumor burden or spread. Since the timeline does not provide direct evidence of infection (e.g., positive cultures, elevated inflammatory markers) or pneumonitis (e.g., post-radiation changes confirmed), but does show multiple cancers, the most consistent diagnosis with the imaging differential is cancer progression.</think> <brief reasoning trace with explanation for label selection> | <label> The patient has documented malignant neoplasms of the lung/bronchus and trachea (C34.80 and C33), supported by CT imaging (CPT4 codes). The nonspecific lung abnormality (R91.8) lacks specific evidence pointing to infection or pneumonitis (no lab results, symptoms, or treatment details). Given the known malignancies and imaging findings, cancer progression is the most consistent diagnosis. <answer>Cancer Progression</answer></think>

no_image: [INCORRECT] <think>First, I need to analyze the patient timeline carefully to understand their medical history and current conditions. The patient has several diagnoses listed: - Hypothyroidism (E03.9) – managed with levothyroxine. - Abdominal aortic aneurysm (I71.4) – non-ruptured. - Malignant neoplasm of overlapping sites of unspecified bronchus and lung (C34.80) – this suggests possible lung or airway malignancy. - Malignant neoplasm of trachea (C33) – indicates another primary or secondary cancer site involving the trachea. - Other nonspecific abnormal lung findings (R91.8) – could relate to infection, inflammation, or malignancy. - Imaging studies: CT scans of thorax and abdomen/pelvis with contrast were performed. - Case management visit noted. The question asks about the imaging differential diagnosis based on these data. The options include pneumonitis, infection, both, cancer progression, or neither. Key points from the timeline: - The presence of two malignant neoplasms (lung/bronchus and trachea) strongly suggests advanced or multifocal cancer. - The CT scans (thorax and abdomen/pelvis) would have been used to evaluate these malignancies. - The nonspecific lung abnormality (R91.8) might indicate an inflammatory process like pneumonitis or infection, but it's not definitive. - There is no mention of biopsy results, lab markers (like white blood cell count), or symptoms such as fever or cough that would support infection. - Pneumonitis can occur due to radiation therapy or chemotherapy, but there's no indication here of recent treatments mentioned. - Cancer progression is plausible given the known malignancies and the imaging findings likely showing tumor burden or spread. Since the timeline does not provide direct evidence of infection (e.g., positive cultures, elevated inflammatory markers) or pneumonitis (e.g., post-radiation changes confirmed), but does show multiple cancers, the most consistent diagnosis with the imaging differential is cancer progression.</think> <brief reasoning trace with explanation for label selection> | <label> The patient has documented malignant neoplasms of the lung/bronchus and trachea (C34.80 and C33), supported by CT imaging (CPT4 codes). The nonspecific lung abnormality (R91.8) lacks specific evidence pointing to infection or pneumonitis (no lab results, symptoms, or treatment details). Given the known malignancies and imaging findings, cancer progression is the most consistent diagnosis. <answer>Cancer Progression</answer></think>

sagittal_all_image: [INCORRECT] <think>First, I need to analyze the patient timeline carefully to understand their medical history and current conditions. The patient has several diagnoses listed: - Hypothyroidism (E03.9) – managed with levothyroxine. - Abdominal aortic aneurysm (I71.4) – non-ruptured. - Malignant neoplasm of overlapping sites of unspecified bronchus and lung (C34.80) – this suggests possible lung or airway malignancy. - Malignant neoplasm of trachea (C33) – indicates another primary or secondary cancer site involving the trachea. - Other nonspecific abnormal lung findings (R91.8) – could relate to infection, inflammation, or malignancy. - Imaging studies: CT scans of thorax and abdomen/pelvis with contrast were performed. - Case management visit noted. The question asks about the imaging differential diagnosis based on these data. The options include pneumonitis, infection, both, cancer progression, or neither. Key points from the timeline: - The presence of two malignant neoplasms (lung/bronchus and trachea) strongly suggests advanced or multifocal cancer. - The CT scans (thorax and abdomen/pelvis) would have been used to evaluate these malignancies. - The nonspecific lung abnormality (R91.8) might indicate an inflammatory process like pneumonitis or infection, but it's not definitive. - There is no mention of biopsy results, lab markers (like white blood cell count), or symptoms such as fever or cough that would support infection. - Pneumonitis can occur due to radiation therapy or chemotherapy, but there's no indication here of recent treatments mentioned. - Cancer progression is plausible given the known malignancies and the imaging findings likely showing tumor burden or spread. Since the timeline does not provide direct evidence of infection (e.g., positive cultures, elevated inflammatory markers) or pneumonitis (e.g., post-radiation changes confirmed), but does show multiple cancers, the most consistent diagnosis with the imaging differential is cancer progression.</think> <brief reasoning trace with explanation for label selection> | <label> The patient has documented malignant neoplasms of the lung/bronchus and trachea (C34.80 and

C33), supported by CT imaging (CPT4 codes). The nonspecific lung abnormality (R91.8) lacks specific evidence pointing to infection or pneumonitis (no lab results, symptoms, or treatment details). Given the known malignancies and imaging findings, cancer progression is the most consistent diagnosis. <answer>Cancer Progression</answer></think>

Model: Qwen/Qwen3-VL-8B-Instruct

all_image: [INCORRECT] The patient has a documented history of Stage IIIA small cell neuroendocrine carcinoma of the left upper lobe, treated with concurrent chemotherapy and radiation, followed by WBRT for metastatic disease. The imaging findings from 2018 (CT and MRI) show multiple lesions: subcentimeter liver nodules, osseous metastases in the lumbar spine, and a pancreatic lesion. The radiology reports explicitly state that the liver nodules are "consistent with hepatic metastases" and the vertebral lesions are "concerning for osseous metastases." This is in the context of a known cancer history and prior treatment for metastatic disease. The imaging findings are not described as inflammatory, infectious, or post-radiation changes alone, but rather as metastatic disease. Therefore, the most consistent diagnosis with the imaging differential is Cancer Progression. <brief reasoning trace with explanation for label selection> | Cancer Progression

axial_all_image: [INCORRECT] The imaging findings and clinical history are consistent with cancer progression. The patient has a history of small cell lung cancer (SCLC) with prior chemotherapy and radiation. The CT scan from 09/04/2018 showed "two subcentimeter hypodense liver lesions" which were later confirmed by MRI on 09/14/2018 as "hepatic metastases." Additionally, the MRI identified "enhancing osseous metastases at L2 vertebral body" and a "questionable lesion in the T11 vertebral body." These findings represent metastatic disease progression, which is the most consistent diagnosis with the imaging differential. The timeline shows a clear progression from initial imaging to confirmation of metastases, and the clinical management plan (ordering PET/CT) is directed at further evaluating cancer spread. There is no evidence of infection or pneumonitis; the imaging findings are consistent with metastatic disease, not inflammatory or infectious processes. | Cancer Progression

no_image: [INCORRECT] The patient has a documented diagnosis of "Malignant neoplasm of overlapping sites of unspecified bronchus and lung" (C34.80) and "Malignant neoplasm of trachea" (C33) — both are cancer diagnoses. The imaging codes (CPT4/71260, 74177) indicate thoracic and abdominal/pelvic CT scans were performed, which are consistent with surveillance or staging for cancer. There is no evidence of infection or pneumonitis in the timeline (no fever, no infectious code, no clinical description of infection or inflammation). The presence of cancer diagnoses and imaging findings in the context of known malignancy strongly supports "Cancer Progression" as the most consistent diagnosis with the imaging differential. | Cancer Progression

sagittal_all_image: [INCORRECT] The patient has a documented history of Stage IIIA small cell neuroendocrine carcinoma of the left upper lobe, treated with concurrent chemotherapy and radiation, followed by WBRT for metastatic disease. The imaging findings from September 2018 show new or enlarging lesions: two subcentimeter hypodense liver lesions (later confirmed as metastases on MRI), and enhancing osseous metastases in the L2 and T11 vertebral bodies. The CT and MRI reports explicitly describe these findings as metastatic disease. The differential diagnosis options include "Pneumonitis", "Infection", "Both", "Cancer Progression", or "Neither". Given the patient's known cancer history and the imaging findings of new metastatic lesions, the most consistent diagnosis is "Cancer Progression". The imaging findings are not consistent with infection or pneumonitis, as there is no evidence of inflammatory or infectious processes described in the reports. Cancer Progression | Cancer Progression

Model: google/medgemma-1.5-4b-it

all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient has a history of malignant neoplasm of the bronchus and lung (C34.80) and trachea (C33), indicating a history of cancer. The CT scan (CPT4/71260) shows findings in the lung field (R91.8). While the specific findings are not detailed, the presence of a known malignancy makes cancer progression a possibility. However, the timeline does not explicitly state any recent cancer-related events, treatment, or worsening symptoms. The CT scan was performed on 2022-09-01, and the specific findings are not described in the provided timeline. Without further information about the nature of the findings on the CT scan or any recent clinical events, it is difficult to definitively distinguish between cancer progression and other possibilities like infection or pneumonitis. The timeline only mentions the initial diagnosis and a CT scan, but not the results or subsequent events. Therefore, based solely on the provided timeline, it is not possible to definitively conclude cancer progression. The timeline does not provide evidence for infection or pneumonitis. Thus, the most appropriate label is "Neither". <label>Neither</label>

axial_all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient has a history of malignant neoplasm of the bronchus and lung (C34.80) and trachea (C33), indicating a history of cancer. The CT scan (CPT4/71260) shows findings in the lung field (R91.8). While the specific findings are not detailed, the presence of a known malignancy makes cancer progression a possibility. However, the timeline does not explicitly state any recent cancer-related events, treatment, or worsening symptoms. The CT scan was performed on 2022-09-01, and the specific findings are not described in the provided timeline. Without further information about the nature of the findings on the CT scan or any recent clinical events, it is difficult to definitively distinguish between cancer progression and other possibilities like infection or pneumonitis. The timeline only mentions the initial diagnosis and a CT scan, but not the results or subsequent events. Therefore, based solely on the provided timeline, it is not possible to definitively conclude cancer progression. The timeline does not provide evidence for infection or pneumonitis. Thus, the most appropriate label is "Neither". <label>Neither</label>

no_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient has a history of malignant neoplasm of the bronchus and lung (C34.80) and trachea (C33), indicating a history of cancer. The CT scan (CPT4/71260) shows findings in the lung field (R91.8). While the specific findings are not detailed, the presence of a known malignancy makes cancer progression a possibility. However, the timeline does not explicitly state any recent cancer-related events, treatment, or worsening symptoms. The CT scan was performed on 2022-09-01, and the specific findings are not described in the provided timeline. Without further information about the nature of the findings on the CT scan or any recent clinical events, it is difficult to definitively distinguish between cancer progression and other possibilities like infection or pneumonitis. The timeline only mentions the initial diagnosis and a CT scan, but not the results or subsequent events. Therefore, based solely on the provided timeline, it is not possible to definitively conclude cancer progression. The timeline does not provide evidence for infection or pneumonitis. Thus, the most appropriate label is "Neither". <label>Neither</label>

sagittal_all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient has a history of malignant neoplasm of the bronchus and lung (C34.80) and trachea (C33), indicating a history of cancer. The CT scan (CPT4/71260) shows findings in the lung field (R91.8). While the specific findings are not detailed, the presence of a known malignancy makes cancer progression a possibility. However, the timeline does not explicitly state any recent cancer-related events, treatment, or worsening symptoms. The CT scan was performed on 2022-09-01, and the specific findings are not described in the provided timeline. Without further information about the nature of the findings on the CT scan or any recent clinical events, it is difficult to definitively distinguish between cancer progression and other possibilities like infection or pneumonitis. The timeline only mentions the initial diagnosis and a CT scan, but not the results or subsequent events. Therefore, based solely on the provided timeline, it is not

possible to definitively conclude cancer progression. The timeline does not provide evidence for infection or pneumonitis. Thus, the most appropriate label is "Neither". <label>Neither</label>

Model: google/medgemma-4b-it

all_image: [INCORRECT] The patient has a history of lung cancer (C34.80, C33) and a nonspecific abnormal finding of the lung field (R91.8). The CT scan shows a mass in the lung. The patient is also being treated with levothyroxine. The imaging differential includes pneumonitis and infection. Given the history of lung cancer, the imaging findings, and the treatment, cancer progression is the most likely diagnosis. <The patient has a history of lung cancer and imaging findings of a mass in the lung. The patient is being treated with levothyroxine. The imaging differential includes pneumonitis and infection. Given the history of lung cancer, the imaging findings, and the treatment, cancer progression is the most likely diagnosis.> | Cancer Progression

axial_all_image: [INCORRECT] The patient has a history of lung cancer (C34.80, C33) and a nonspecific abnormal finding of the lung field (R91.8). The CT scan shows a mass in the lung. The patient is also being treated with levothyroxine. The imaging differential includes pneumonitis and infection. Given the history of lung cancer, the imaging findings, and the treatment, cancer progression is the most likely diagnosis. <The patient has a history of lung cancer and imaging findings of a mass in the lung. The patient is being treated with levothyroxine. The imaging differential includes pneumonitis and infection. Given the history of lung cancer, the imaging findings, and the treatment, cancer progression is the most likely diagnosis.> | Cancer Progression

no_image: [INCORRECT] The patient has a history of lung cancer (C34.80, C33) and a nonspecific abnormal finding of the lung field (R91.8). The CT scan shows a mass in the lung. The patient is also being treated with levothyroxine. The imaging differential includes pneumonitis and infection. Given the history of lung cancer, the imaging findings, and the treatment, cancer progression is the most likely diagnosis. <The patient has a history of lung cancer and imaging findings of a mass in the lung. The patient is being treated with levothyroxine. The imaging differential includes pneumonitis and infection. Given the history of lung cancer, the imaging findings, and the treatment, cancer progression is the most likely diagnosis.> | Cancer Progression

sagittal_all_image: [INCORRECT] The patient has a history of lung cancer (C34.80, C33) and a nonspecific abnormal finding of the lung field (R91.8). The CT scan shows a mass in the lung. The patient is also being treated with levothyroxine. The imaging differential includes pneumonitis and infection. Given the history of lung cancer, the imaging findings, and the treatment, cancer progression is the most likely diagnosis. <The patient has a history of lung cancer and imaging findings of a mass in the lung. The patient is being treated with levothyroxine. The imaging differential includes pneumonitis and infection. Given the history of lung cancer, the imaging findings, and the treatment, cancer progression is the most likely diagnosis.> | Cancer Progression

Subtask: pneumonitis_infection_discussed

Question:

Based on the provided medical history, cancer evidence, and imaging findings, is there sufficient diagnostic ambiguity to warrant presentation of this case at a thoracic tumor board to discuss differentiation between radiation pneumonitis and infection?

Correct Answer: 0.0

Model: OpenGVLab/InternVL3_5-8B

all_image: [INCORRECT] <no evidence of cancer or relevant imaging findings to suggest differentiation between radiation pneumonitis and infection> | No

axial_all_image: [INCORRECT] <no evidence of cancer or relevant imaging findings to suggest differentiation between radiation pneumonitis and infection> | No

no_image: [INCORRECT] <no evidence of cancer or relevant imaging findings to suggest differentiation between radiation pneumonitis and infection> | No

sagittal_all_image: [INCORRECT] <no evidence of cancer or relevant imaging findings to suggest differentiation between radiation pneumonitis and infection> | No

Model: OctoMed/OctoMed-7B

all_image: [INCORRECT] <think>First, I need to determine if there is enough diagnostic ambiguity to present this case at a thoracic tumor board to differentiate between radiation pneumonitis and infection based on the patient's timeline. Looking through the patient timeline, the key points related to potential radiation pneumonitis or infection would involve any treatments or conditions affecting the lungs or chest area post-radiation therapy. However, the timeline does not mention any radiation therapy to the thoracic region. Instead, the focus seems to be on gastrointestinal biopsies, breast imaging, and various lab tests related to autoimmune or other conditions. The patient had a colon biopsy showing normal mast cells, and subsequent breast imaging and procedures indicating no malignancy. There are mentions of mast cell activation disorder, eating disorders, depression, thyroiditis, and past tobacco use, but none directly relate to lung issues or recent radiation therapy. Radiation pneumonitis typically occurs after radiation therapy targeting the chest, especially for cancers like lung or breast cancer. Since there is no indication of prior radiation therapy in the timeline, nor any symptoms or imaging findings suggestive of lung inflammation or infection, there is no basis to suspect radiation pneumonitis here. Moreover, the imaging studies mentioned include abdominal and breast imaging, and later brain MRI, but no thoracic imaging or notes about lung pathology. Without any evidence of lung involvement or recent radiation, differentiating between radiation pneumonitis and infection isn't warranted. Therefore, since the timeline lacks any evidence supporting radiation pneumonitis or infection in the thoracic region, there is no diagnostic ambiguity requiring a tumor board discussion.</think> <brief reasoning trace with explanation for label selection> | <label> - The patient's timeline includes gastrointestinal biopsies, breast imaging, and various lab tests but no evidence of thoracic radiation therapy, lung pathology, or symptoms indicative of radiation pneumonitis or infection. - Radiation pneumonitis requires prior thoracic radiation, which is absent here. - No imaging or clinical findings related to the lungs or chest are documented. - Therefore, there is no diagnostic ambiguity necessitating a tumor board discussion. </think> <answer>Yes</answer> | <answer>No</answer> <answer>No</answer> **Final Answer:** \boxed{No}

axial_all_image: [INCORRECT] <think>First, I need to determine if there is enough diagnostic ambiguity to present this case at a thoracic tumor board to differentiate between radiation pneumonitis and infection based on the patient's timeline. Looking through the patient timeline, the key points related to potential radiation pneumonitis or infection would involve any treatments or conditions affecting the lungs or chest area post-radiation therapy. However, the timeline does not mention any radiation therapy to the thoracic region. Instead, the focus seems to be on gastrointestinal biopsies, breast imaging, and various lab tests related to autoimmune or other conditions. The patient had a colon biopsy showing normal mast cells, and subsequent breast imaging and procedures indicating no malignancy. There are mentions of mast cell activation disorder, eating disorders, depression, thyroiditis, and past tobacco use, but none directly relate to lung issues or recent radiation therapy. Radiation pneumonitis typically occurs after radiation therapy targeting the chest, especially for cancers like lung or breast cancer. Since there is no indication of prior radiation therapy in the timeline, nor any symptoms or imaging findings suggestive of lung inflammation or infection, there is no basis to suspect radiation pneumonitis here. Moreover, the imaging studies mentioned include abdominal and breast imaging, and later brain MRI, but no thoracic imaging or notes about lung pathology. Without any evidence of lung involvement or recent radiation, differentiating between radiation pneumonitis and infection isn't warranted. Therefore, since the timeline lacks any evidence supporting radiation pneumonitis or infection in the thoracic region, there is no diagnostic ambiguity requiring a tumor board discussion.</think> <brief reasoning trace with explanation for label selection> | <label> - The patient's timeline includes gastrointestinal biopsies, breast imaging, and various lab tests but no evidence of thoracic radiation therapy, lung pathology, or symptoms indicative of radiation pneumonitis or infection. - Radiation pneumonitis requires prior thoracic radiation, which is absent here. - No imaging or clinical findings related to the lungs or chest are documented. - Therefore, there is no diagnostic ambiguity necessitating a tumor board discussion. </think> <answer>Yes</answer> | <answer>No</answer> <answer>No</answer> **Final Answer:** \boxed{No}

no_image: [INCORRECT] <think>First, I need to determine if there is enough diagnostic ambiguity to present this case at a thoracic tumor board to differentiate between radiation pneumonitis and infection based on the patient's timeline. Looking through the patient timeline, the key points related to potential radiation pneumonitis or infection would involve any treatments or conditions affecting the lungs or chest area post-radiation therapy. However, the timeline does not mention any radiation therapy to the thoracic region. Instead, the focus seems to be on gastrointestinal biopsies, breast imaging, and various lab tests related to autoimmune or other conditions. The patient had a colon biopsy showing normal mast cells, and subsequent breast imaging and procedures indicating no malignancy. There are mentions of mast cell activation disorder, eating disorders, depression, thyroiditis, and past tobacco use, but none directly relate to lung issues or recent radiation therapy. Radiation pneumonitis typically occurs after radiation therapy targeting the chest, especially for cancers like lung or breast cancer. Since there is no indication of prior radiation therapy in the timeline, nor any symptoms or imaging findings suggestive of lung inflammation or infection, there is no basis to suspect radiation pneumonitis here. Moreover, the imaging studies mentioned include abdominal and breast imaging, and later brain MRI, but no thoracic imaging or notes about lung pathology. Without any evidence of lung involvement or recent radiation, differentiating between radiation pneumonitis and infection isn't warranted. Therefore, since the timeline lacks any evidence supporting radiation pneumonitis or infection in the thoracic region, there is no diagnostic ambiguity requiring a tumor board discussion.</think> <brief reasoning trace with explanation for label selection> | <label> - The patient's timeline includes gastrointestinal biopsies, breast imaging, and various lab tests but no evidence of thoracic radiation therapy, lung pathology, or symptoms indicative of radiation pneumonitis or infection. - Radiation pneumonitis requires prior thoracic radiation, which is absent here. - No imaging or clinical findings related to the lungs or chest are documented. - Therefore, there is no diagnostic ambiguity necessitating a tumor

board discussion. </think> <answer>Yes</answer> | <answer>No</answer> <answer>No</answer> **Final Answer:** \boxed{No}

sagittal_all_image: [INCORRECT] <think>First, I need to determine if there is enough diagnostic ambiguity to present this case at a thoracic tumor board to differentiate between radiation pneumonitis and infection based on the patient's timeline. Looking through the patient timeline, the key points related to potential radiation pneumonitis or infection would involve any treatments or conditions affecting the lungs or chest area post-radiation therapy. However, the timeline does not mention any radiation therapy to the thoracic region. Instead, the focus seems to be on gastrointestinal biopsies, breast imaging, and various lab tests related to autoimmune or other conditions. The patient had a colon biopsy showing normal mast cells, and subsequent breast imaging and procedures indicating no malignancy. There are mentions of mast cell activation disorder, eating disorders, depression, thyroiditis, and past tobacco use, but none directly relate to lung issues or recent radiation therapy. Radiation pneumonitis typically occurs after radiation therapy targeting the chest, especially for cancers like lung or breast cancer. Since there is no indication of prior radiation therapy in the timeline, nor any symptoms or imaging findings suggestive of lung inflammation or infection, there is no basis to suspect radiation pneumonitis here. Moreover, the imaging studies mentioned include abdominal and breast imaging, and later brain MRI, but no thoracic imaging or notes about lung pathology. Without any evidence of lung involvement or recent radiation, differentiating between radiation pneumonitis and infection isn't warranted. Therefore, since the timeline lacks any evidence supporting radiation pneumonitis or infection in the thoracic region, there is no diagnostic ambiguity requiring a tumor board discussion.</think> <brief reasoning trace with explanation for label selection> | <label> - The patient's timeline includes gastrointestinal biopsies, breast imaging, and various lab tests but no evidence of thoracic radiation therapy, lung pathology, or symptoms indicative of radiation pneumonitis or infection. - Radiation pneumonitis requires prior thoracic radiation, which is absent here. - No imaging or clinical findings related to the lungs or chest are documented. - Therefore, there is no diagnostic ambiguity necessitating a tumor board discussion. </think> <answer>Yes</answer> | <answer>No</answer> <answer>No</answer> **Final Answer:** \boxed{No}

Model: Qwen/Qwen3-VL-8B-Instruct

all_image: [INCORRECT] No | There is no evidence in the provided timeline of thoracic imaging, radiation therapy history, or clinical findings suggestive of radiation pneumonitis or infection requiring differentiation. The timeline includes breast imaging, systemic evaluations, and GI biopsies, but no thoracic imaging, radiation exposure, or pulmonary symptoms are documented. Therefore, diagnostic ambiguity for thoracic tumor board discussion does not exist.

no_image: [INCORRECT] No | There is no evidence in the provided timeline of thoracic imaging, radiation therapy history, or clinical findings suggestive of radiation pneumonitis or infection requiring differentiation. The timeline includes breast imaging, systemic evaluations, and GI biopsies, but no thoracic imaging, radiation exposure, or pulmonary symptoms are documented. Therefore, diagnostic ambiguity for thoracic tumor board discussion does not exist.

Model: google/medgemma-1.5-4b-it

no_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label> The patient presents with a history of right nipple soreness, which was investigated with a diagnostic mammogram and ultrasound in January 2014. Both imaging studies were negative, showing no suspicious masses or findings. The patient subsequently underwent a biopsy of a lesion, which was also negative. The patient has a history of mast cell activation disorder, but

the biopsy findings were not consistent with malignancy. The patient has also had a history of headaches and other conditions, but no evidence of malignancy is present in the provided timeline. The imaging findings are negative for malignancy. No | Yes

Model: google/medgemma-4b-it

Subtask: progression_assessment_discussed

Question:

Based on the provided medical history and cancer evidence, is there ambiguity in radiologic interpretation that warrants presentation of this case at a thoracic tumor board for progression assessment?

Correct Answer: 0.0

Model: OpenGVLab/InternVL3_5-8B

all_image: [INCORRECT] <no radiologic evidence of progression or ambiguity in interpretation> | No

axial_all_image: [INCORRECT] <no radiologic evidence of progression or ambiguity in interpretation> | No

no_image: [INCORRECT] <no radiologic evidence of progression or ambiguity in interpretation> | No

sagittal_all_image: [INCORRECT] <no radiologic evidence of progression or ambiguity in interpretation> | No

Model: OctoMed/OctoMed-7B

all_image: [INCORRECT] <think>First, I need to determine if there is any ambiguity in the radiologic interpretation that would warrant presenting this case at a thoracic tumor board for progression assessment based on the provided timeline. Looking through the patient timeline, the key points related to imaging and cancer evaluation are: - On 2013-12-02, there is a mammogram and ultrasound report indicating no suspicious findings. The BI-RADS scores are both 1 (negative), suggesting no malignancy detected. - The patient had a biopsy of various gastrointestinal sites (duodenum, stomach, distal esophagus, proximal esophagus, ileum, colon) with pathology showing mostly mild chronic gastritis, squamous epithelium without dysplasia, and no significant abnormalities except for some mast cell presence noted in the colon biopsy. - There is mention of a mast cell activation disorder workup later, but no indication of malignancy in these biopsies. - The patient underwent multiple imaging studies including MRI of the brain (normal), MR angiogram (no aneurysm), and other tests like CT scans and lab work, all without evidence of cancer progression. - The clinical history includes a history of mast cell activation disorder, but no mention of cancer recurrence or progression. - The most recent imaging (MRI brain in 2014-09-05) also showed no abnormalities. There is no indication in the timeline of any new or progressive thoracic malignancy. The biopsies do not indicate cancer, and imaging does not show suspicious lesions. Therefore, there is no ambiguity in the radiologic interpretation that would require a tumor board review for progression assessment.</think> <brief reasoning trace with explanation for label selection> | <label> The patient's timeline demonstrates no evidence of thoracic malignancy or progression. Multiple imaging studies (mammogram, ultrasound, MRI) consistently reported negative findings (BI-RADS 1) with no suspicious masses or abnormalities. Biopsy results from gastrointestinal sites revealed benign conditions (e.g., mild gastritis, no dysplasia) without malignancy. There is no indication of cancer recurrence or progression requiring multidisciplinary tumor board review. Thus, there is no ambiguity in radiologic interpretation necessitating a thoracic tumor board presentation. </brief reasoning trace with explanation for label selection> | <label> <answer>No</answer></think>

axial_all_image: [INCORRECT] <think>First, I need to determine if there is any ambiguity in the radiologic interpretation that would warrant presenting this case at a thoracic tumor board for progression assessment based on the provided timeline. Looking through the patient timeline, the key points related to imaging and cancer evaluation are: - On 2013-12-02, there is a mammogram and ultrasound report indicating no suspicious findings. The BI-RADS scores are both 1 (negative), suggesting no malignancy detected. - The patient had a biopsy of various gastrointestinal sites (duodenum, stomach, distal esophagus, proximal esophagus, ileum, colon) with pathology showing mostly mild chronic gastritis, squamous epithelium without dysplasia, and no significant abnormalities except for some mast cell presence noted in the colon biopsy. - There is mention of a mast cell activation disorder workup later, but no indication of malignancy in these biopsies. - The patient underwent multiple imaging studies including MRI of the brain (normal), MR angiogram (no aneurysm), and other tests like CT scans and lab work, all without evidence of cancer progression. - The clinical history includes a history of mast cell activation disorder, but no mention of cancer recurrence or progression. - The most recent imaging (MRI brain in 2014-09-05) also showed no abnormalities. There is no indication in the timeline of any new or progressive thoracic malignancy. The biopsies do not indicate cancer, and imaging does not show suspicious lesions. Therefore, there is no ambiguity in the radiologic interpretation that would require a tumor board review for progression assessment.</think> <brief reasoning trace with explanation for label selection> | <label> The patient's timeline demonstrates no evidence of thoracic malignancy or progression. Multiple imaging studies (mammogram, ultrasound, MRI) consistently reported negative findings (BI-RADS 1) with no suspicious masses or abnormalities. Biopsy results from gastrointestinal sites revealed benign conditions (e.g., mild gastritis, no dysplasia) without malignancy. There is no indication of cancer recurrence or progression requiring multidisciplinary tumor board review. Thus, there is no ambiguity in radiologic interpretation necessitating a thoracic tumor board presentation. </brief reasoning trace with explanation for label selection> | <label> <answer>No</answer></think>

no_image: [INCORRECT] <think>First, I need to determine if there is any ambiguity in the radiologic interpretation that would warrant presenting this case at a thoracic tumor board for progression assessment based on the provided timeline. Looking through the patient timeline, the key points related to imaging and cancer evaluation are: - On 2013-12-02, there is a mammogram and ultrasound report indicating no suspicious findings. The BI-RADS scores are both 1 (negative), suggesting no malignancy detected. - The patient had a biopsy of various gastrointestinal sites (duodenum, stomach, distal esophagus, proximal esophagus, ileum, colon) with pathology showing mostly mild chronic gastritis, squamous epithelium without dysplasia, and no significant abnormalities except for some mast cell presence noted in the colon biopsy. - There is mention of a mast cell activation disorder workup later, but no indication of malignancy in these biopsies. - The patient underwent multiple imaging studies including MRI of the brain (normal), MR angiogram (no aneurysm), and other tests like CT scans and lab work, all without evidence of cancer progression. - The clinical history includes a history of mast cell activation disorder, but no mention of cancer recurrence or progression. - The most recent imaging (MRI brain in 2014-09-05) also showed no abnormalities. There is no indication in the timeline of any new or progressive thoracic malignancy. The biopsies do not indicate cancer, and imaging does not show suspicious lesions. Therefore, there is no ambiguity in the radiologic interpretation that would require a tumor board review for progression assessment.</think> <brief reasoning trace with explanation for label selection> | <label> The patient's timeline demonstrates no evidence of thoracic malignancy or progression. Multiple imaging studies (mammogram, ultrasound, MRI) consistently reported negative findings (BI-RADS 1) with no suspicious masses or abnormalities. Biopsy results from gastrointestinal sites revealed benign conditions (e.g., mild gastritis, no dysplasia) without malignancy. There is no indication of cancer recurrence or progression requiring multidisciplinary tumor board review. Thus, there is no ambiguity in radiologic interpretation necessitating a thoracic tumor board presentation.

</brief reasoning trace with explanation for label selection> | <label>
<answer>No</answer></think>

sagittal_all_image: [INCORRECT] <think>First, I need to determine if there is any ambiguity in the radiologic interpretation that would warrant presenting this case at a thoracic tumor board for progression assessment based on the provided timeline. Looking through the patient timeline, the key points related to imaging and cancer evaluation are: - On 2013-12-02, there is a mammogram and ultrasound report indicating no suspicious findings. The BI-RADS scores are both 1 (negative), suggesting no malignancy detected. - The patient had a biopsy of various gastrointestinal sites (duodenum, stomach, distal esophagus, proximal esophagus, ileum, colon) with pathology showing mostly mild chronic gastritis, squamous epithelium without dysplasia, and no significant abnormalities except for some mast cell presence noted in the colon biopsy. - There is mention of a mast cell activation disorder workup later, but no indication of malignancy in these biopsies. - The patient underwent multiple imaging studies including MRI of the brain (normal), MR angiogram (no aneurysm), and other tests like CT scans and lab work, all without evidence of cancer progression. - The clinical history includes a history of mast cell activation disorder, but no mention of cancer recurrence or progression. - The most recent imaging (MRI brain in 2014-09-05) also showed no abnormalities. There is no indication in the timeline of any new or progressive thoracic malignancy. The biopsies do not indicate cancer, and imaging does not show suspicious lesions. Therefore, there is no ambiguity in the radiologic interpretation that would require a tumor board review for progression assessment.</think>
<brief reasoning trace with explanation for label selection> | <label> The patient's timeline demonstrates no evidence of thoracic malignancy or progression. Multiple imaging studies (mammogram, ultrasound, MRI) consistently reported negative findings (BI-RADS 1) with no suspicious masses or abnormalities. Biopsy results from gastrointestinal sites revealed benign conditions (e.g., mild gastritis, no dysplasia) without malignancy. There is no indication of cancer recurrence or progression requiring multidisciplinary tumor board review. Thus, there is no ambiguity in radiologic interpretation necessitating a thoracic tumor board presentation.
</brief reasoning trace with explanation for label selection> | <label>
<answer>No</answer></think>

Model: Qwen/Qwen3-VL-8B-Instruct

all_image: [INCORRECT] No | There is no evidence of ambiguous radiologic interpretation or progression assessment required. All imaging (mammogram, ultrasound, MRI brain) were interpreted as negative (BI-RADS 1) and no suspicious findings were reported. No thoracic tumor board referral is warranted based on the provided timeline.

axial_all_image: [INCORRECT] No | There is no evidence of ambiguous radiologic interpretation or progression assessment required. All imaging (mammogram, ultrasound, MRI brain) were interpreted as negative (BI-RADS 1) and no suspicious findings were reported. No thoracic tumor board referral is warranted based on the provided timeline.

no_image: [INCORRECT] No | There is no evidence of ambiguous radiologic interpretation or progression assessment required. All imaging (mammogram, ultrasound, MRI brain) were interpreted as negative (BI-RADS 1) and no suspicious findings were reported. No thoracic tumor board referral is warranted based on the provided timeline.

sagittal_all_image: [INCORRECT] No | There is no evidence of ambiguous radiologic interpretation or progression assessment required. All imaging (mammogram, ultrasound, MRI brain) were interpreted as negative (BI-RADS 1) and no suspicious findings were reported. No thoracic tumor board referral is warranted based on the provided timeline.

Model: google/medgemma-1.5-4b-it

all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient underwent a diagnostic mammogram and ultrasound on January 13, 2014, which were negative for malignancy. The mammogram showed heterogeneously dense breast tissue, no new focal mass, architectural distortion, or suspicious microcalcifications. The ultrasound revealed no discrete cystic or solid masses. The impression was BI-RADS 1, Negative for both the right and left breast. The patient reported right nipple soreness, which was noted in the mammogram report and follow-up notes. The follow-up mammogram on January 13, 2014, was also negative. The patient was recommended to return for routine annual screening mammogram and follow-up for nipple soreness. There is no evidence of progression or suspicious findings in the provided timeline. No | No

axial_all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> |
<label> The patient underwent a diagnostic mammogram and ultrasound on January 13, 2014, which were negative for malignancy. The mammogram showed heterogeneously dense breast tissue, no new focal mass, architectural distortion, or suspicious microcalcifications. The ultrasound revealed no discrete cystic or solid masses. The impression was BI-RADS 1, Negative for both the right and left breast. The patient reported right nipple soreness, which was noted in the mammogram report and follow-up notes. The follow-up mammogram on January 13, 2014, was also negative. The patient was recommended to return for routine annual screening mammogram and follow-up for nipple soreness. There is no evidence of progression or suspicious findings in the provided timeline. No | No

no_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient underwent a diagnostic mammogram and ultrasound on January 13, 2014, which were negative for malignancy. The mammogram showed heterogeneously dense breast tissue, no new focal mass, architectural distortion, or suspicious microcalcifications. The ultrasound revealed no discrete cystic or solid masses. The impression was BI-RADS 1, Negative for both the right and left breast. The patient reported right nipple soreness, which was noted in the mammogram report and follow-up notes. The follow-up mammogram on January 13, 2014, was also negative. The patient was recommended to return for routine annual screening mammogram and follow-up for nipple soreness. There is no evidence of progression or suspicious findings in the provided timeline. No | No

sagittal_all_image: [INCORRECT] <brief reasoning trace with explanation for label selection>
<label> The patient underwent a diagnostic mammogram and ultrasound on January 13, 2014, which were negative for malignancy. The mammogram showed heterogeneously dense breast tissue, no new focal mass, architectural distortion, or suspicious microcalcifications. The ultrasound revealed no discrete cystic or solid masses. The impression was BI-RADS 1, Negative for both the right and left breast. The patient reported right nipple soreness, which was noted in the mammogram report and follow-up notes. The follow-up mammogram on January 13, 2014, was also negative. The patient was recommended to return for routine annual screening mammogram and follow-up for nipple soreness. There is no evidence of progression or suspicious findings in the provided timeline. No | No

Model: google/medgemma-4b-it

axial_all_image: [INCORRECT] The patient has a history of small intestinal biopsy showing no significant abnormality in 2013. The patient had a mammogram in 2014 showing no malignancy. The patient had a colonoscopy in 2013 with a random colon biopsy showing a relatively normal distribution of CD117+ mast cells. The patient has a history of smoking. The patient has a

breast cancer diagnosis. The patient has a history of a prior breast cancer diagnosis. The patient has a history of a prior breast cancer diagnosis. The patient has a history of a prior breast cancer diagnosis. The patient has a history of a prior breast cancer diagnosis. The patient has a history

Subtask: radiation_oligoprogression_answer

Question:

TODO: Based on the provided medical history, cancer evidence, and imaging findings, is focal radiotherapy recommended for the management of this patient's oligopressive disease?

Correct Answer: -1

Model: OpenGVLab/InternVL3_5-8B

all_image: [INCORRECT] <The patient's medical history indicates a secondary malignant neoplasm of the brain (ICD10CM/C79.31) and other parts of the nervous system (ICD10CM/C79.49). The timeline mentions a planned cyberknife treatment to the resection bed and a follow-up with thoracic oncology. Given the context of oligopression and the planned focal radiotherapy (cyberknife), focal radiotherapy is appropriate for managing oligopression in this scenario.> | Yes

axial_all_image: [INCORRECT] <The patient's medical history indicates a secondary malignant neoplasm of the brain (ICD10CM/C79.31) and other parts of the nervous system (ICD10CM/C79.49). The timeline mentions a planned cyberknife treatment to the resection bed and a follow-up with thoracic oncology. Given the context of oligopression and the planned focal radiotherapy (cyberknife), focal radiotherapy is appropriate for managing oligopression in this scenario.> | Yes

no_image: [INCORRECT] <The patient's medical history indicates a secondary malignant neoplasm of the brain (ICD10CM/C79.31) and other parts of the nervous system (ICD10CM/C79.49). The timeline mentions a planned cyberknife treatment to the resection bed and a follow-up with thoracic oncology. Given the context of oligopression and the planned focal radiotherapy (cyberknife), it is appropriate to consider focal radiotherapy as a management option.> | <Yes>

sagittal_all_image: [INCORRECT] <The patient's medical history indicates a secondary malignant neoplasm of the brain (ICD10CM/C79.31) and other parts of the nervous system (ICD10CM/C79.49). The timeline mentions a planned cyberknife treatment to the resection bed and a follow-up with thoracic oncology. Given the context of oligopression and the planned cyberknife treatment, focal radiotherapy is appropriate for managing the progression.> | Yes

Model: OctoMed/OctoMed-7B

all_image: [INCORRECT] <think>First, I need to understand what the question is asking. It wants to know if focal radiotherapy is appropriate for managing oligopression based on the patient's timeline. Oligopression refers to limited sites of disease progression after initial treatment. Looking at the patient timeline: - Multiple entries on 2021-06-07 with ICD10CM/D49.6, which is "Neoplasm of unspecified behavior of brain." This suggests the patient had a primary brain tumor diagnosis around that date. - On 2021-06-11, there are two entries: C79.31 ("Secondary malignant neoplasm of brain") and C79.49 ("Secondary malignant neoplasm of other parts of nervous system"). These indicate metastatic spread to the brain and possibly other parts of the nervous system. The patient underwent surgery (as implied by the note about not submerging the head until 4 weeks post-surgery) and then was scheduled for CyberKnife treatment to the resection bed. CyberKnife is a form of stereotactic radiosurgery, which is a type of focal radiotherapy used for precise targeting of small tumors or residual

disease. The key point here is that the patient has secondary malignant neoplasms (metastases) involving the brain and possibly other nervous system areas. Oligoprogression would mean progression at only a few sites. However, the presence of multiple secondary sites (brain and other nervous system) suggests more than just oligoprogression; it might be multifocal or widespread metastatic disease. Focal radiotherapy like CyberKnife is typically considered when there are limited metastatic lesions (oligometastatic disease), usually 1-3 or up to 5 lesions depending on guidelines, to control those sites without systemic therapy. But here, the patient already has secondary involvement in both brain and other nervous system regions, indicating a higher burden of disease beyond oligoprogression. Moreover, the timeline shows that the patient is still undergoing treatment (CyberKnife scheduled) but now has new codes indicating secondary malignancies. This implies that the initial treatment may have failed, leading to metastases. In such cases, focal radiotherapy alone may not be sufficient because the disease is disseminated rather than localized. Additionally, the decision to use focal radiotherapy depends on factors like tumor biology, prior treatments, margins, etc., but none of these details are provided except through the codes. Since the patient has secondary malignant neoplasms, especially in multiple sites, focal therapy is less likely to be appropriate as monotherapy. Therefore, based on the timeline showing secondary metastases in multiple sites, focal radiotherapy is not indicated for oligoprogression here.</think> <brief reasoning trace with explanation for label selection> | <label> The patient's timeline indicates secondary malignant neoplasms involving the brain (C79.31) and other parts of the nervous system (C79.49), suggesting metastatic spread beyond a single site. Focal radiotherapy is generally reserved for oligometastatic disease (limited number of metastases). Here, the presence of multiple secondary sites implies a broader disease burden, making focal radiotherapy inappropriate as standalone management. Additionally, the patient is still undergoing treatment (CyberKnife scheduled), but the progression to secondary malignancies signals advanced disease requiring systemic evaluation rather than localized intervention. Thus, focal radiotherapy is not suitable for oligoprogression in this context.</think> \boxed{No}

axial_all_image: [INCORRECT] <think>First, I need to understand what the question is asking. It wants to know if focal radiotherapy is appropriate for managing oligoprogression based on the patient's timeline. Oligoprogression refers to limited sites of disease progression after initial treatment. Looking at the patient timeline: - Multiple entries on 2021-06-07 with ICD10CM/D49.6, which is "Neoplasm of unspecified behavior of brain." This suggests the patient had a primary brain tumor diagnosis around that date. - On 2021-06-11, there are two entries: C79.31 ("Secondary malignant neoplasm of brain") and C79.49 ("Secondary malignant neoplasm of other parts of nervous system"). These indicate metastatic spread to the brain and possibly other parts of the nervous system. The patient underwent surgery (as implied by the note about not submerging the head until 4 weeks post-surgery) and then was scheduled for CyberKnife treatment to the resection bed. CyberKnife is a form of stereotactic radiosurgery, which is a type of focal radiotherapy used for precise targeting of small tumors or residual disease. The key point here is that the patient has secondary malignant neoplasms (metastases) involving the brain and possibly other nervous system areas. Oligoprogression would mean progression at only a few sites. However, the presence of multiple secondary sites (brain and other nervous system) suggests more than just oligoprogression; it might be multifocal or widespread metastatic disease. Focal radiotherapy like CyberKnife is typically considered when there are limited metastatic lesions (oligometastatic disease), usually 1-3 or up to 5 lesions depending on guidelines, to control those sites without systemic therapy. But here, the patient already has secondary involvement in both brain and other nervous system regions, indicating a higher burden of disease beyond oligoprogression. Moreover, the timeline shows that the patient is still undergoing treatment (CyberKnife scheduled) but now has new codes indicating secondary malignancies. This implies that the initial treatment may have failed, leading to metastases. In such cases, focal radiotherapy alone may not be sufficient because the disease is disseminated rather than localized. Additionally, the decision to use focal

radiotherapy depends on factors like tumor biology, prior treatments, margins, etc., but none of these details are provided except through the codes. Since the patient has secondary malignant neoplasms, especially in multiple sites, focal therapy is less likely to be appropriate as monotherapy. Therefore, based on the timeline showing secondary metastases in multiple sites, focal radiotherapy is not indicated for oligoprogression here.</think> <brief reasoning trace with explanation for label selection> | <label> The patient's timeline indicates secondary malignant neoplasms involving the brain (C79.31) and other parts of the nervous system (C79.49), suggesting metastatic spread beyond a single site. Focal radiotherapy is generally reserved for oligometastatic disease (limited number of metastases). Here, the presence of multiple secondary sites implies a broader disease burden, making focal radiotherapy inappropriate as standalone management. Additionally, the patient is still undergoing treatment (CyberKnife scheduled), but the progression to secondary malignancies signals advanced disease requiring systemic evaluation rather than localized intervention. Thus, focal radiotherapy is not suitable for oligoprogression in this context.</think> \boxed{No}

no_image: **[INCORRECT]** <think>First, I need to understand what the question is asking. It wants to know if focal radiotherapy is appropriate for managing oligoprogression based on the patient's timeline. Oligoprogression refers to limited sites of disease progression after initial treatment. Looking at the patient timeline: - Multiple entries on 2021-06-07 with ICD10CM/D49.6, which is "Neoplasm of unspecified behavior of brain." This suggests the patient had a primary brain tumor diagnosis around that date. - On 2021-06-11, there are two entries: C79.31 ("Secondary malignant neoplasm of brain") and C79.49 ("Secondary malignant neoplasm of other parts of nervous system"). These indicate metastatic spread to the brain and possibly other parts of the nervous system. The patient underwent surgery (as implied by the note about not submerging the head until 4 weeks post-surgery) and then was scheduled for CyberKnife treatment to the resection bed. CyberKnife is a form of stereotactic radiosurgery, which is a type of focal radiotherapy used for precise targeting of small tumors or residual disease. The key point here is that the patient has secondary malignant neoplasms (metastases) involving the brain and possibly other nervous system areas. Oligoprogression would mean progression at only a few sites. However, the presence of multiple secondary sites (brain and other nervous system) suggests more than just oligoprogression; it might be multifocal or widespread metastatic disease. Focal radiotherapy like CyberKnife is typically considered when there are limited metastatic lesions (oligometastatic disease), usually 1-3 or up to 5 lesions depending on guidelines, to control those sites without systemic therapy. But here, the patient already has secondary involvement in both brain and other nervous system regions, indicating a higher burden of disease beyond oligoprogression. Moreover, the timeline shows that the patient is still undergoing treatment (CyberKnife scheduled) but now has new codes indicating secondary malignancies. This implies that the initial treatment may have failed, leading to metastases. In such cases, focal radiotherapy alone may not be sufficient because the disease is disseminated rather than localized. Additionally, the decision to use focal radiotherapy depends on factors like tumor biology, prior treatments, margins, etc., but none of these details are provided except through the codes. Since the patient has secondary malignant neoplasms, especially in multiple sites, focal therapy is less likely to be appropriate as monotherapy. Therefore, based on the timeline showing secondary metastases in multiple sites, focal radiotherapy is not indicated for oligoprogression here.</think> <brief reasoning trace with explanation for label selection> | <label> The patient's timeline indicates secondary malignant neoplasms involving the brain (C79.31) and other parts of the nervous system (C79.49), suggesting metastatic spread beyond a single site. Focal radiotherapy is generally reserved for oligometastatic disease (limited number of metastases). Here, the presence of multiple secondary sites implies a broader disease burden, making focal radiotherapy inappropriate as standalone management. Additionally, the patient is still undergoing treatment (CyberKnife scheduled), but the progression to secondary malignancies signals advanced disease requiring systemic evaluation rather than localized intervention. Thus, focal radiotherapy is not suitable

for oligoprogression in this context.</think> \boxed{No}

sagittal_all_image: [INCORRECT] <think>First, I need to understand what the question is asking. It wants to know if focal radiotherapy is appropriate for managing oligoprogression based on the patient's timeline. Oligoprogression refers to limited sites of disease progression after initial treatment. Looking at the patient timeline: - Multiple entries on 2021-06-07 with ICD10CM/D49.6, which is "Neoplasm of unspecified behavior of brain." This suggests the patient had a primary brain tumor diagnosis around that date. - On 2021-06-11, there are two entries: C79.31 ("Secondary malignant neoplasm of brain") and C79.49 ("Secondary malignant neoplasm of other parts of nervous system"). These indicate metastatic spread to the brain and possibly other parts of the nervous system. The patient underwent surgery (as implied by the note about not submerging the head until 4 weeks post-surgery) and then was scheduled for CyberKnife treatment to the resection bed. CyberKnife is a form of stereotactic radiosurgery, which is a type of focal radiotherapy used for precise targeting of small tumors or residual disease. The key point here is that the patient has secondary malignant neoplasms (metastases) involving the brain and possibly other nervous system areas. Oligoprogression would mean progression at only a few sites. However, the presence of multiple secondary sites (brain and other nervous system) suggests more than just oligoprogression; it might be multifocal or widespread metastatic disease. Focal radiotherapy like CyberKnife is typically considered when there are limited metastatic lesions (oligometastatic disease), usually 1-3 or up to 5 lesions depending on guidelines, to control those sites without systemic therapy. But here, the patient already has secondary involvement in both brain and other nervous system regions, indicating a higher burden of disease beyond oligoprogression. Moreover, the timeline shows that the patient is still undergoing treatment (CyberKnife scheduled) but now has new codes indicating secondary malignancies. This implies that the initial treatment may have failed, leading to metastases. In such cases, focal radiotherapy alone may not be sufficient because the disease is disseminated rather than localized. Additionally, the decision to use focal radiotherapy depends on factors like tumor biology, prior treatments, margins, etc., but none of these details are provided except through the codes. Since the patient has secondary malignant neoplasms, especially in multiple sites, focal therapy is less likely to be appropriate as monotherapy. Therefore, based on the timeline showing secondary metastases in multiple sites, focal radiotherapy is not indicated for oligoprogression here.</think> <brief reasoning trace with explanation for label selection> | <label> The patient's timeline indicates secondary malignant neoplasms involving the brain (C79.31) and other parts of the nervous system (C79.49), suggesting metastatic spread beyond a single site. Focal radiotherapy is generally reserved for oligometastatic disease (limited number of metastases). Here, the presence of multiple secondary sites implies a broader disease burden, making focal radiotherapy inappropriate as standalone management. Additionally, the patient is still undergoing treatment (CyberKnife scheduled), but the progression to secondary malignancies signals advanced disease requiring systemic evaluation rather than localized intervention. Thus, focal radiotherapy is not suitable for oligoprogression in this context.</think> \boxed{No}

Model: Qwen/Qwen3-VL-8B-Instruct

all_image: [INCORRECT] The patient has Stage IVB metastatic thymoma with documented progressive disease. The most recent imaging (06/28/2018 and 09/17/2018) shows continued progression, specifically in the right upper hemithorax and right pleura. The clinical note from 07/27/2018 explicitly states that the patient is on a treatment break and that scans have shown "slowly progressive disease" without significant clinical manifestations. The plan is to repeat scans in 2 months and to consider starting everolimus if there is continued symptomatic worsening. The patient's goals of care are centered on quality of life, and she is currently feeling well. There is no mention of oligoprogression or any specific lesion that would be a candidate for focal radiotherapy. The management plan is to continue observation and consider systemic

therapy if symptoms worsen, not focal radiotherapy. No | The patient's disease is described as "slowly progressive" and the management plan is to continue on a treatment break with repeat scans in 2 months, not to initiate focal radiotherapy. There is no evidence of oligoprogression or indication for focal therapy in the provided timeline.

axial_all_image: [INCORRECT] The patient has Stage IVB metastatic thymic neoplasm with documented progressive disease. The most recent imaging (06/28/2018 and 09/17/2018) shows continued progression, specifically in the right upper hemithorax with enlarged pleural nodules and effusion. The clinical note from 07/27/2018 explicitly states that the patient is on a treatment break and that scans have shown "slowly progressive disease" without significant clinical manifestations. The plan is to repeat scans in 2 months and to consider starting everolimus if there is continued symptomatic worsening. The patient's goals of care are centered on quality of life, and she is currently feeling well. There is no evidence of oligoprogression (a limited number of sites of progression) that would warrant focal radiotherapy. The disease is widespread and progressive, and the management plan is to continue observation and consider systemic therapy if symptoms worsen. No | The patient's disease is broadly progressive, not oligopressive, and the management plan is observation with potential future systemic therapy, not focal radiotherapy.

no_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label> The patient's timeline indicates a history of brain neoplasm (D49.6) and secondary malignant neoplasms of the brain and nervous system (C79.31, C79.49), with documented post-operative care and planned cyberknife treatment. However, there is no evidence of oligoprogression (i.e., limited progression of disease after initial treatment) or imaging findings confirming focal progression requiring radiotherapy. The timeline does not mention any documented progression, imaging evidence of new or enlarging lesions, or clinical indication for focal radiotherapy. Therefore, based on the provided evidence, focal radiotherapy is not indicated for oligoprogression. | No

sagittal_all_image: [INCORRECT] The patient has Stage IVB metastatic thymoma with documented progressive disease. The most recent imaging (09/17/2018) shows continued progression, specifically "continued progression of right-sided predominant pleural metastatic disease" and "new apparent soft tissue defect involving chest wall". The clinical note from 07/27/2018 explicitly states that the patient is on a treatment break and that scans have continued to show "slowly progressive disease". The plan is to repeat scans in 2 months and to consider starting everolimus if there is continued symptomatic worsening. There is no mention of oligoprogression or any indication that focal radiotherapy is being considered or would be appropriate. The patient's goals of care are focused on quality of life, and she is currently stable on a treatment break. No | The patient's disease is described as "slowly progressive" and "continued progression" without evidence of oligoprogression. The plan is to continue on a treatment break and repeat scans, not to initiate focal radiotherapy.

Model: google/medgemma-1.5-4b-it

all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label> The patient timeline shows the patient was diagnosed with a neoplasm of unspecified behavior of the brain in June 2021, followed by a secondary malignant neoplasm of the brain and other parts of the nervous system in June 2021. There is no mention of imaging findings, treatment history, or documented evidence of oligoprogression. Therefore, based solely on the provided timeline, it is not possible to determine if focal radiotherapy is appropriate. No

axial_all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label> The patient timeline shows the patient was diagnosed with a neoplasm of unspecified behavior of the brain in June 2021, followed by a secondary malignant neoplasm of the brain and other parts of the nervous system in June 2021. There is no mention of imaging findings,

treatment history, or documented evidence of oligoprogression. Therefore, based solely on the provided timeline, it is not possible to determine if focal radiotherapy is appropriate. No

no_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient timeline shows the patient was diagnosed with a neoplasm of unspecified behavior of the brain in June 2021, followed by a secondary malignant neoplasm of the brain and other parts of the nervous system in June 2021. There is no mention of imaging findings, treatment history, or documented evidence of oligoprogression. Therefore, based solely on the provided timeline, it is not possible to determine if focal radiotherapy is appropriate. No

sagittal_all_image: [INCORRECT] <brief reasoning trace with explanation for label selection> | <label>
The patient timeline shows the patient was diagnosed with a neoplasm of unspecified behavior of the brain in June 2021, followed by a secondary malignant neoplasm of the brain and other parts of the nervous system in June 2021. There is no mention of imaging findings, treatment history, or documented evidence of oligoprogression. Therefore, based solely on the provided timeline, it is not possible to determine if focal radiotherapy is appropriate. No

Model: google/medgemma-4b-it

all_image: [INCORRECT] The patient has a history of brain neoplasms, including secondary malignant neoplasms. The patient is currently being followed by thoracic oncology. The patient has had cyberknife treatment. The patient is experiencing oligoprogression. Focal radiotherapy is an option for oligoprogression. Yes

axial_all_image: [INCORRECT] The patient has a history of brain neoplasms, including secondary malignant neoplasms. The patient is currently being followed by thoracic oncology. The patient has had cyberknife treatment. The patient is experiencing oligoprogression. Focal radiotherapy is an option for oligoprogression. Yes

no_image: [INCORRECT] The patient has a history of brain neoplasms, including secondary malignant neoplasms. The patient is currently being followed by thoracic oncology. The patient has had cyberknife treatment. The patient is experiencing oligoprogression. Focal radiotherapy is an option for oligoprogression. Yes

sagittal_all_image: [INCORRECT] The patient has a history of brain neoplasms, including secondary malignant neoplasms. The patient is currently being followed by thoracic oncology. The patient has had cyberknife treatment. The patient is experiencing oligoprogression. Focal radiotherapy is an option for oligoprogression. Yes