Response evaluation and adverse events

Part 1 of 2

Tomoko Kataoka
JCOG(*) Operations Office

The 23rd JCOG Clinical Trial Seminar
10/10/2020

*** Japan Clinical Oncology Group** (https://jcog.jp/en/)

JCOG
Japan Clinical Oncology Group

ATLAS
ASIAN CLINICAL TRIALS NETWORK FOR CANCERS PROJECT

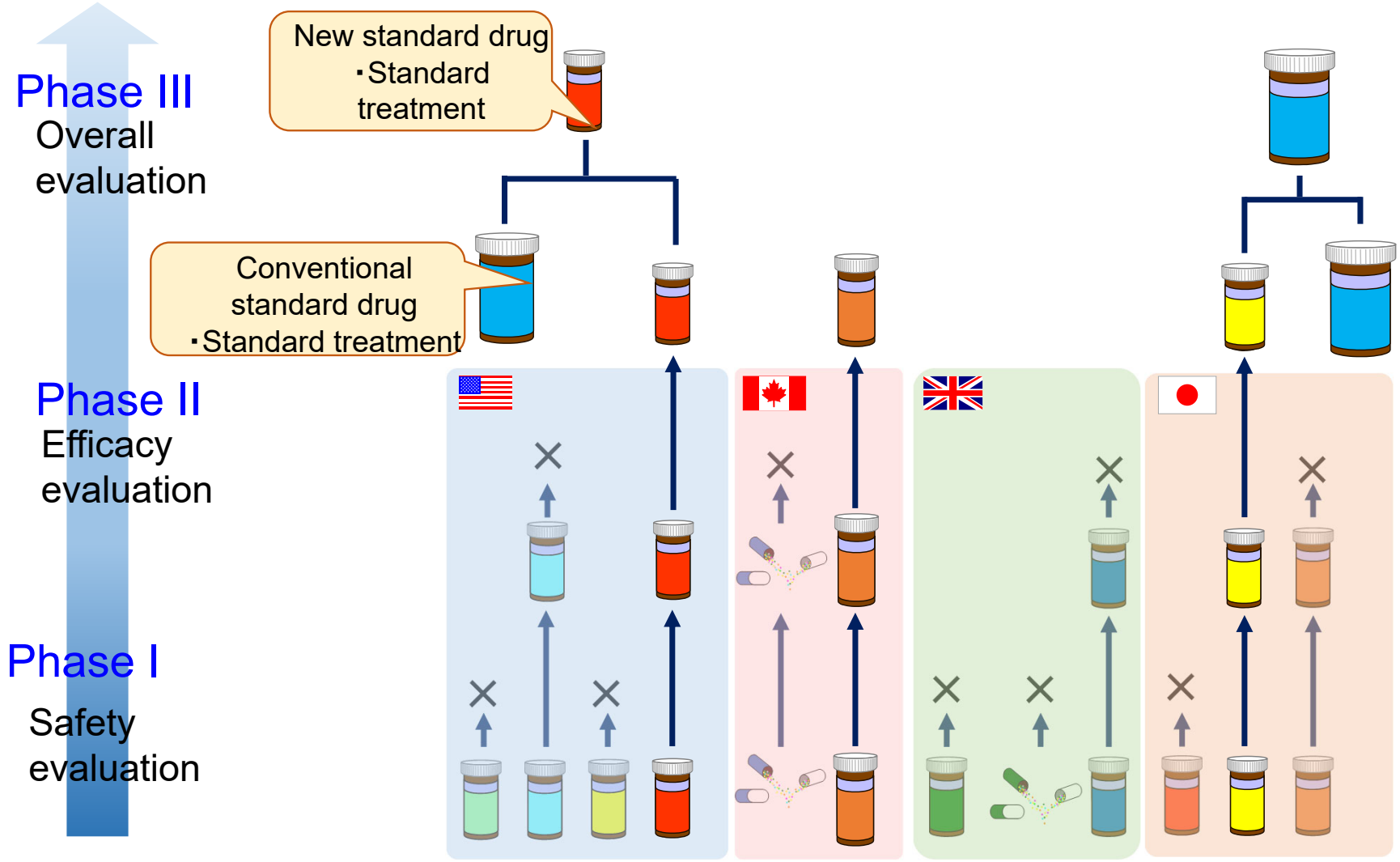
Today's contents

- Response evaluation and
  Response Evaluation Criteria in Solid Tumors (RECIST)
  - Accuracy and precision
  - Basic logic of RECIST
  - Hypothetical example of response evaluation of a lesion

- Adverse events and
  Common Terminology Criteria for Adverse Events (CTCAE)
  - What is an adverse event?
  - Reporting adverse events
  - History and structure of CTCAE
  - Evaluation of adverse events using CTCAE

Why are “common international” standards needed?

National Cancer Center Japan



Why are “common international” standards needed?

National Cancer Center Japan

Phase III
Overall evaluation

Phase II
Efficacy evaluation

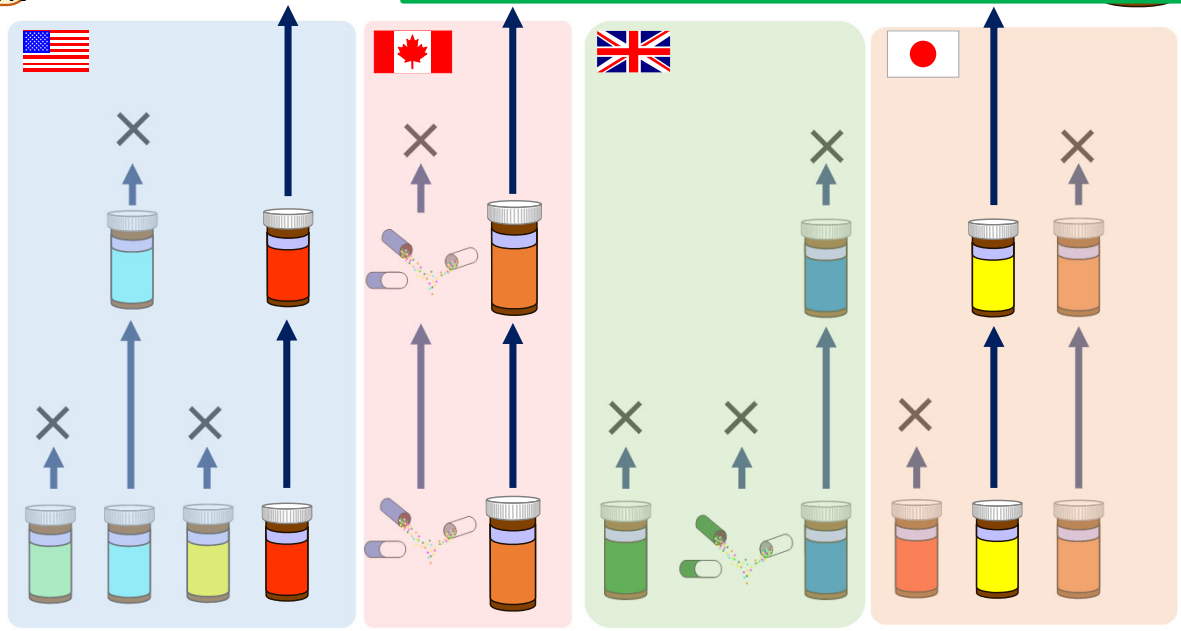
Phase I
Safety evaluation

New standard drug
・Standard treatment

Conventional standard drug
・Standard treatment

Comparison with results from other clinical trials is necessary
- To advance development correctly to the next phase
- To avoid unnecessary duplication of clinical trials

Criteria for comparison across clinical trials
- Efficacy evaluation: RECIST
- Safety evaluation: CTCAE



Today's contents

- Response evaluation and
  Response Evaluation Criteria in Solid Tumors (RECIST)
  - Accuracy and precision
  - Basic logic of RECIST
  - Hypothetical example of response evaluation of a lesion

- Adverse events and
  Common Terminology Criteria for Adverse Events
  - What is an adverse event?
  - Reporting adverse events
  - History and structure of CTCAE
  - Evaluation of adverse events using CTCAE

EUROPEAN JOURNAL OF CANCER 45 (2009) 228–247

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1)

E.A. Eisenhauer[a,*], P. Therasse[b], J. Bogaerts[c], L.H. Schwartz[d], D. Sargent[e], R. Ford[f], J. Dancey[g], S. Arbuck[h], S. Gwyther[i], M. Mooney[g], L. Rubinstein[g], L. Shankar[g], L. Dodd[g], R. Kaplan[j], D. Lacombe[c], J. Verweij[k]

[a]National Cancer Institute of Canada – Clinical Trials Group, 10 Stuart Street, Queen's University, Kingston, ON, Canada
[b]GlaxoSmithKline Biologicals, Rixensart, Belgium
[c]European Organisation for Research and Treatment of Cancer, Data Centre, Brussels, Belgium
[d]Memorial Sloan Kettering Cancer Center, New York, NY, USA
[e]Mayo Clinic, Rochester, MN, USA
[f]RadPharm, Princeton, NJ, USA
[g]Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD, USA
[h]Schering-Plough, Kenilworth, NJ, USA
[i]East Surrey Hospital, Redhill, Surrey, UK
[j]National Cancer Research Network, Leeds, UK
[k]Erasmus University Medical Center, Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received 17 October 2008
Accepted 29 October 2008

Keywords:
Response criteria
Solid tumours
Guidelines

ABSTRACT

Background: Assessment of the change in tumour burden is an important feature of the clinical evaluation of cancer therapeutics: both tumour shrinkage (objective response) and disease progression are useful endpoints in clinical trials. Since RECIST was published in 2000, many investigators, cooperative groups, industry and government authorities have adopted these criteria in the assessment of treatment outcomes. However, a number of questions and issues have arisen which have led to the development of a revised RECIST guideline (version 1.1). Evidence for changes, summarised in separate papers in this special issue, has come from assessment of a large data warehouse (>6500 patients), simulation studies and literature reviews.

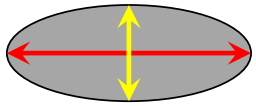
Highlights of revised RECIST 1.1: Major changes include: *Number of lesions to be assessed:* based on evidence from numerous trial databases merged into a data warehouse for analysis purposes, the number of lesions required to assess tumour burden for response determination has been reduced from a maximum of 10 to a maximum of five total (and from five to two per organ, maximum). *Assessment of pathological lymph* nodes is now incorporated: nodes with a short axis of ≥15 mm are considered measurable and assessable as target lesions. The short axis measurement should be included in the sum of lesions in calculation of tumour response. Nodes that shrink to <10 mm short axis are considered normal. *Confirmation of response* is required for trials with response primary endpoint but is no longer required in randomised studies since the control arm serves as appropriate means of interpretation of data. *Disease progression* is clarified in several aspects: in addition to the previous definition of progression in target disease of 20% increase in sum, a 5 mm absolute increase is now required as well to guard against over calling PD when the total sum is very

* *Corresponding author:* Tel.: +1 613 533 6430; fax: +1 613 533 2411.
E-mail address: eeisenhauer@ctg.queensu.ca (E.A. Eisenhauer).
0959-8049/$ - see front matter
doi:10.1016/j.ejca.2008.10.026

Why RECIST were developed

- WHO response evaluation criteria (1981)
  - Many unclear aspects
    - Criteria for selection of lesion to evaluate
    - Calculation of overall response: Per lesion? Per organ? As a whole?
    - Definition of progressive disease (PD): Per lesion? Per organ? As a whole?
  - Additions and modifications made by each research group
  - Evaluated by bi-directional area (major axis × minor axis) for each lesion
- RECIST ver. 1.0 (2000)
  - Response rates in single-arm Phase II trials worldwide are now comparable [standardization]
  - Bi-directional measurement → uni-directional measurement [simplification]
- RECIST ver. 1.1 (2009)
  - Clarification of the definition of pathologic lymphadenopathy as a measurable lesion [standardization]
  - Clarification of the definition of progressive disease (PD) [standardization]
  - Change in the number of target lesions from 10 → 5, etc. [simplification]

Cannot compare between groups



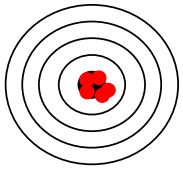
What are RECIST a “measure” for?

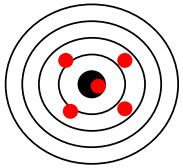
- Index for “Phase III" go / no-go tasks = tool for clinical trials
  - RECIST prioritizes "precision" and "comparability" over "accuracy"
  - "Used to judge whether a drug or regimen shows promising results that warrant continued developmental research”
  - Judgment for whether to proceed to a Phase III trial based on results from a single-arm Phase II trial
- It is not an index for deciding whether to continue or discontinue "treatment"
  - Whether to continue or discontinue treatment should be decided based on clinical / comprehensive judgment in daily clinical practice and in clinical trials and need not be decided with RECIST (in fact, it is inappropriate to do so)

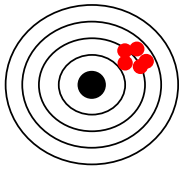
  e.g., A "50% decrease" in tumor cross-section does not necessarily have an essential meaning for individual patients.

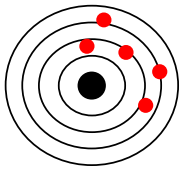
From "New Response Evaluation Criteria in Solid Tumors (RECIST Guidelines) -Japanese translation JCOG version-" Introduction

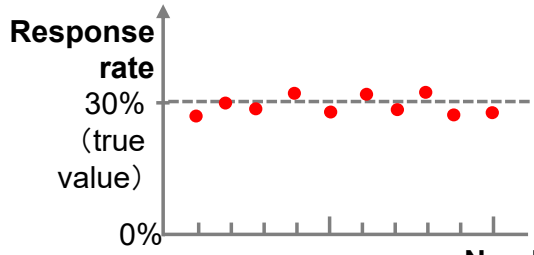
Accuracy vs. Precision

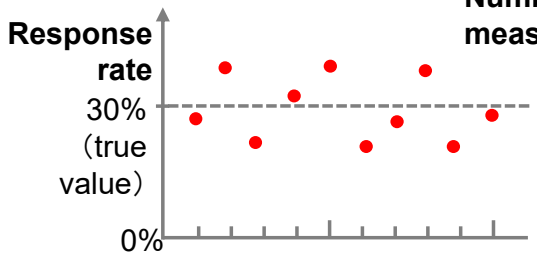


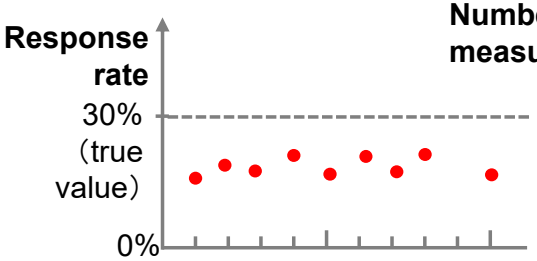


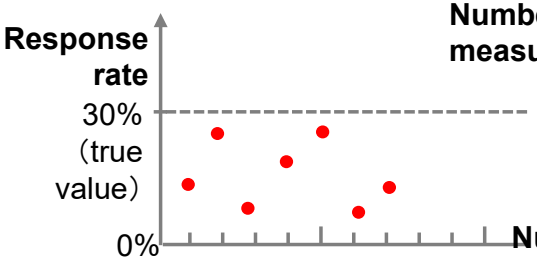


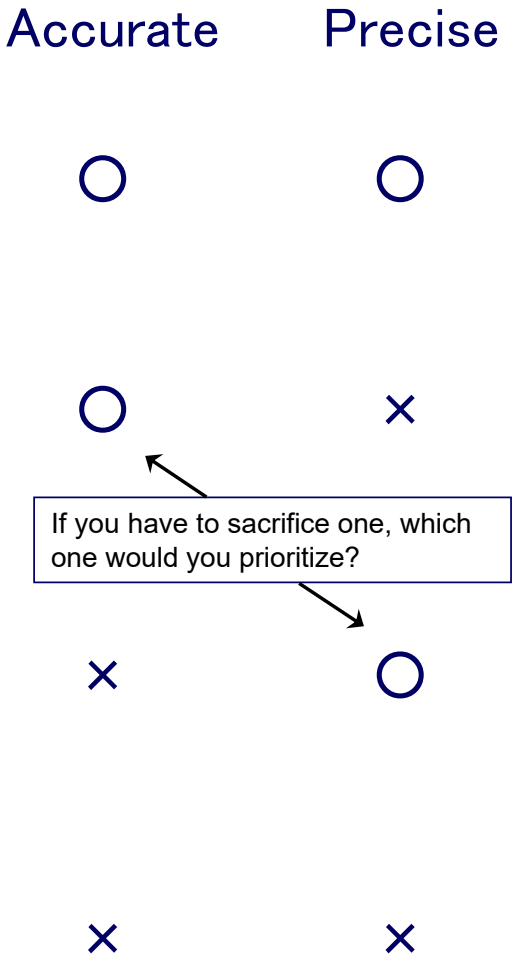






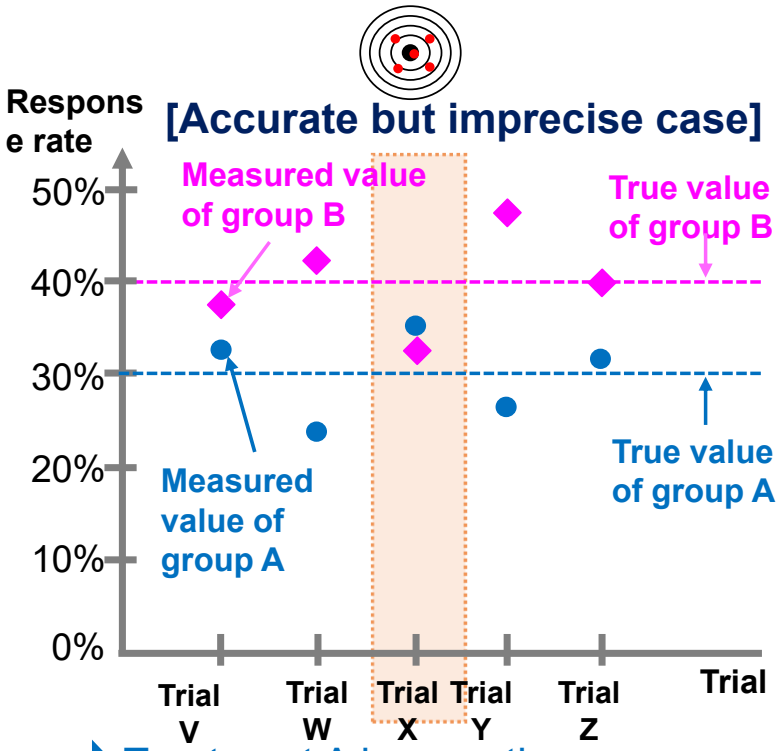




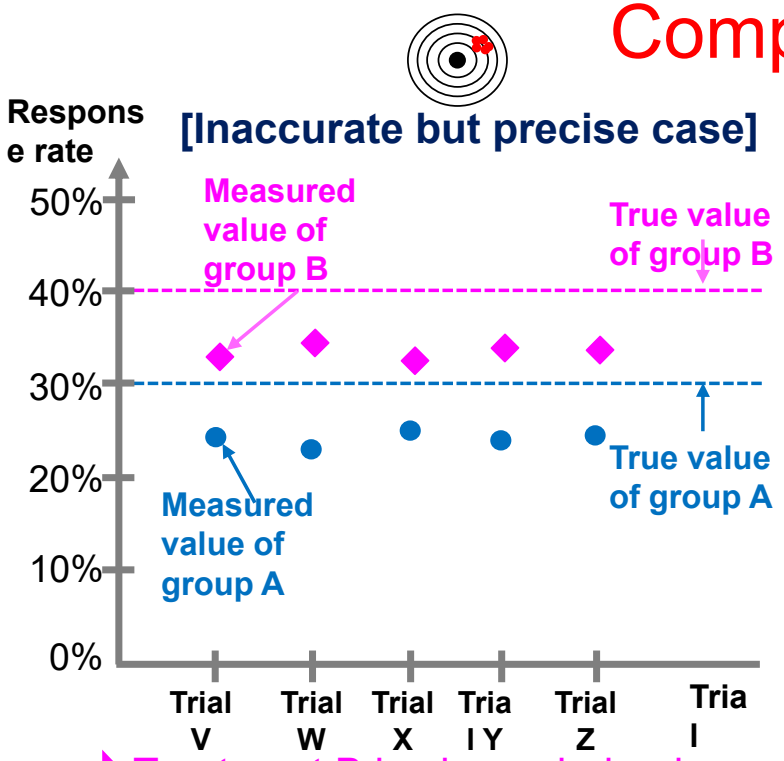


Accuracy vs. Precision

Comparability



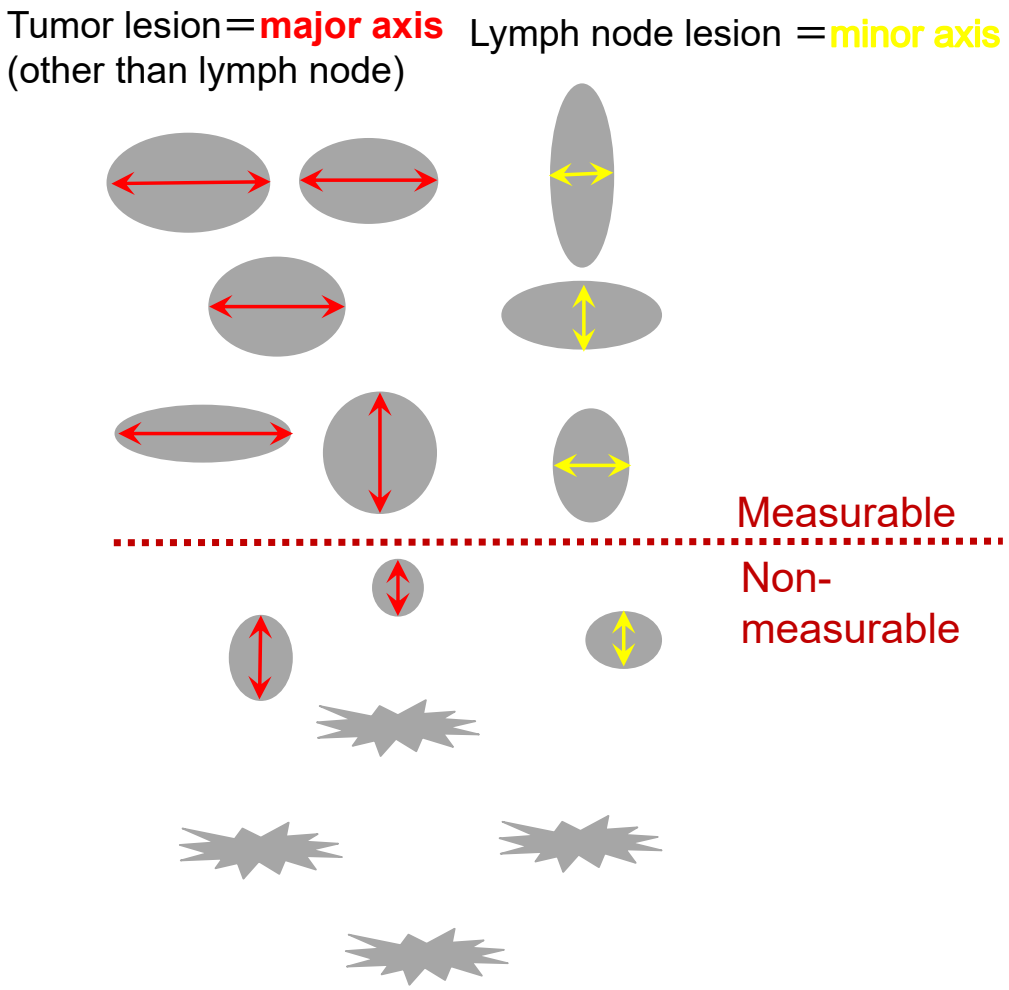
Treatment A is sometimes judged as more promising (error)



Treatment B is always judged as more promising (correct)

- If everyone uses the criteria that emphasize "comparability" (precision), then it is more likely that "more promising treatments" will proceed correctly to Phase III trials
- Underestimation itself is not a problem if underestimated worldwide

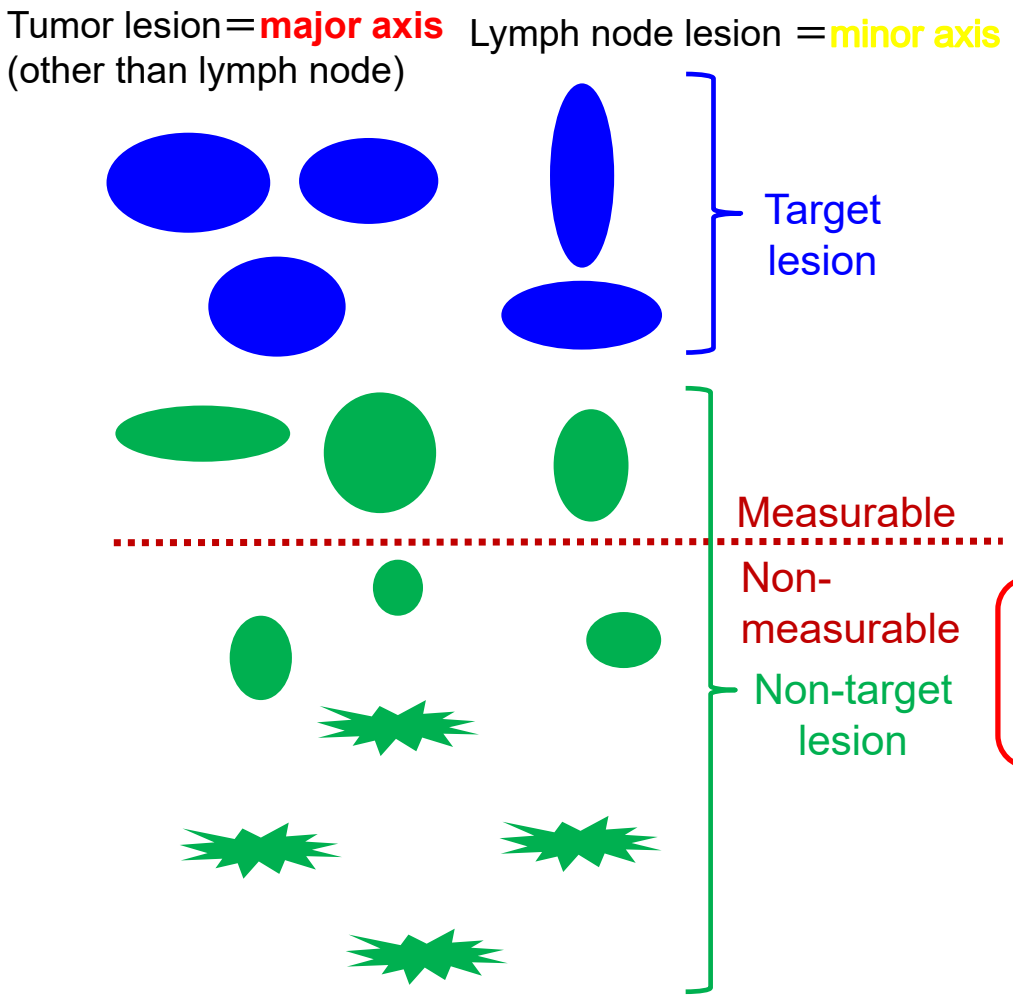
Basic logic of RECIST



① All lesions are classified as “measurable” or “non-measurable”

- Tumor lesion (other than lymph node)：
  Major axis ≥10 mm
- Lymph node lesion：
  Minor axis ≥15 mm

Basic logic of RECIST



① All lesions are classified as “measurable” or “non-measurable”
- Tumor lesion (other than lymph node)：
  Major axis ≥10 mm
- Lymph node lesion：
  Minor axis ≥15 mm

② Select “target lesion” from measurable lesion
- Up to 5 in total (up to 2 per organ)
- All lesions other than target lesion are non-target lesions

Pre-treatment evaluation should be performed as close to the start of treatment as possible (Guideline: within 4 weeks)

③ Conduct response evaluation for each category
- Target lesion effect
- Non-target lesion effect
- New lesion presence

Basic logic of RECIST (continued)

CR: complete response
PR: partial response
SD: stable disease
PD: progressive disease

③ Determine the effect of each course in each category
- Target lesion effect CR, PR, SD, PD
- Non-target lesion effect CR, Non-CR/Non-PD, PD
- New lesion presence None, present

④ Determine "overall response" by combining the effects for each category
- Judge "overall response" for each course

⑤ Determine the "best overall response"
- Determine one "best overall response" from "overall responses" of all courses
- Response rate: the proportion of patients with "best overall response" of PR or better

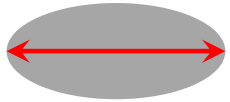
	After 1 course	After 2 courses	After 3 courses	After 4 courses	After 5 courses
Target lesion effect	SD	PR	PR	CR	PD
Non-target lesion effect	non-CR/ non-PD	non-CR/ non-PD	non-CR/ non-PD	CR	CR
New lesion presence	None	None	None	None	Present
Overall response	**SD**	**PR**	**PR**	**CR**	**PD**

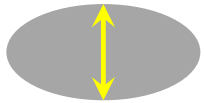
Calculate the "best overall response" among all courses

(1) Classification as “measurable” or “non-measurable”

- “Measurable” lesion
  - Tumor lesion (non-lymph node lesion)
    - Major axis 10 mm or more (CT slice thickness is 5 mm or less)

      For slice thickness > 5 mm, major axis is at least twice the slice thickness
    - Osteolytic bone lesions with soft tissue components can be measurable lesions
  - Lymph node lesion
    - Minor axis 15 mm or more (CT slice thickness is 5 mm or less)





- “Non-measurable” lesion
  - Small lesion
    - Tumor lesion (non-lymph node lesion)
      - Major axis less than 10 mm, but lesion likely present
    - Lymph node lesion
      - Minor axis 10-15 mm

        *Lymph nodes with a minor axis of less than 10 mm are not considered lesions = normal
  - Truly non-measurable lesion
    - Ascites, pleural effusion, pericardial effusion, inflammatory breast cancer, meningeal lesions, etc.

(2) Selection of “target lesions” from measurable lesions

- Target lesion
  - Up to 5 measurable lesions in descending order of diameter
  - Up to two per organ
    - How to count one organ? (not described in RECIST ver. 1.1)

      → Stipulations required for each trial

      <Standard counting method for JCOG trial>
      - Combine organs with left and right sides (e.g., lungs, kidneys) into one organ
      - Treat all the lymph nodes of the body as one organ, regardless of location

        (up to two lymph nodes are included in target lesions per patient)
  - Be sure to include all organs with measurable lesions
  - Easy to measure and reproducible

    (avoid lesions that are difficult to measure even when having a large diameter)
- Non-target lesion
  - Everything “besides” target lesions are non-target lesions

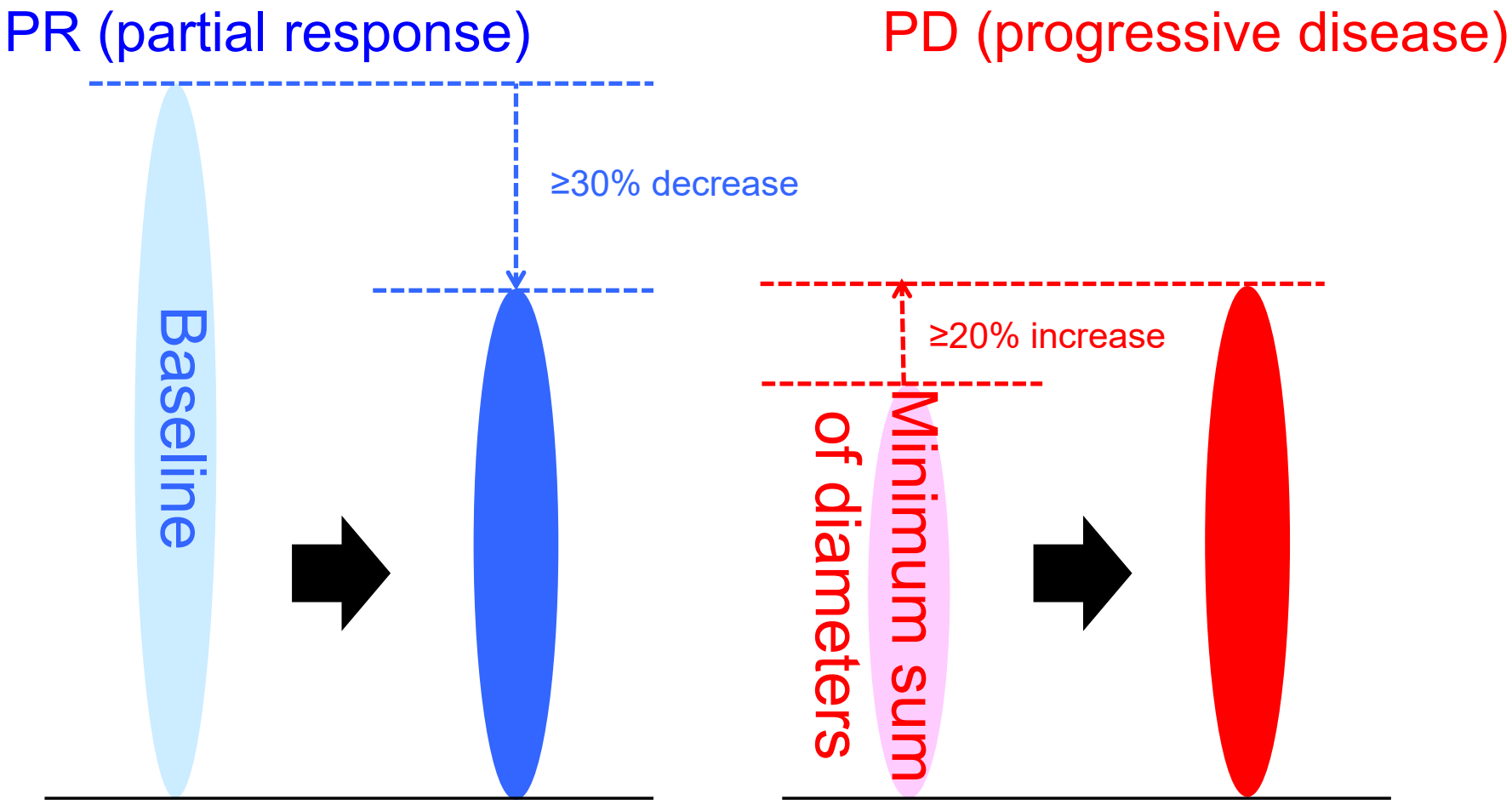
(3) Determining the effect for each category

National Cancer Center Japan

- Target lesion effect: decrease / increase in the of sum of diameters
  - CR (complete response): Disappearance of all tumor lesions
Lymph node lesion has a minor axis of less than 10 mm
  - PR (partial response): Sum of diameters decreased by at least 30% relative to baseline
  - SD (stable disease): Neither CR / PR nor PD
  - PD (progressive disease): Sum of diameters increased by at least 20% relative to the minimum value during the course
- Non-target lesion effect: “disappeared” or “increased”
  - CR (complete response): Disappearance of all tumor lesions
Lymph node lesion has a minor axis of less than 10 mm
Tumor markers below upper limit of reference range
  - Non-CR/non-PD: One or more non-target lesions remaining
Tumor marker exceeds the upper limit of reference range
  - PD (progressive disease): Clear increase in growth of non-target lesion (unequivocal progression)
- New lesion presence: “present” or “absent”

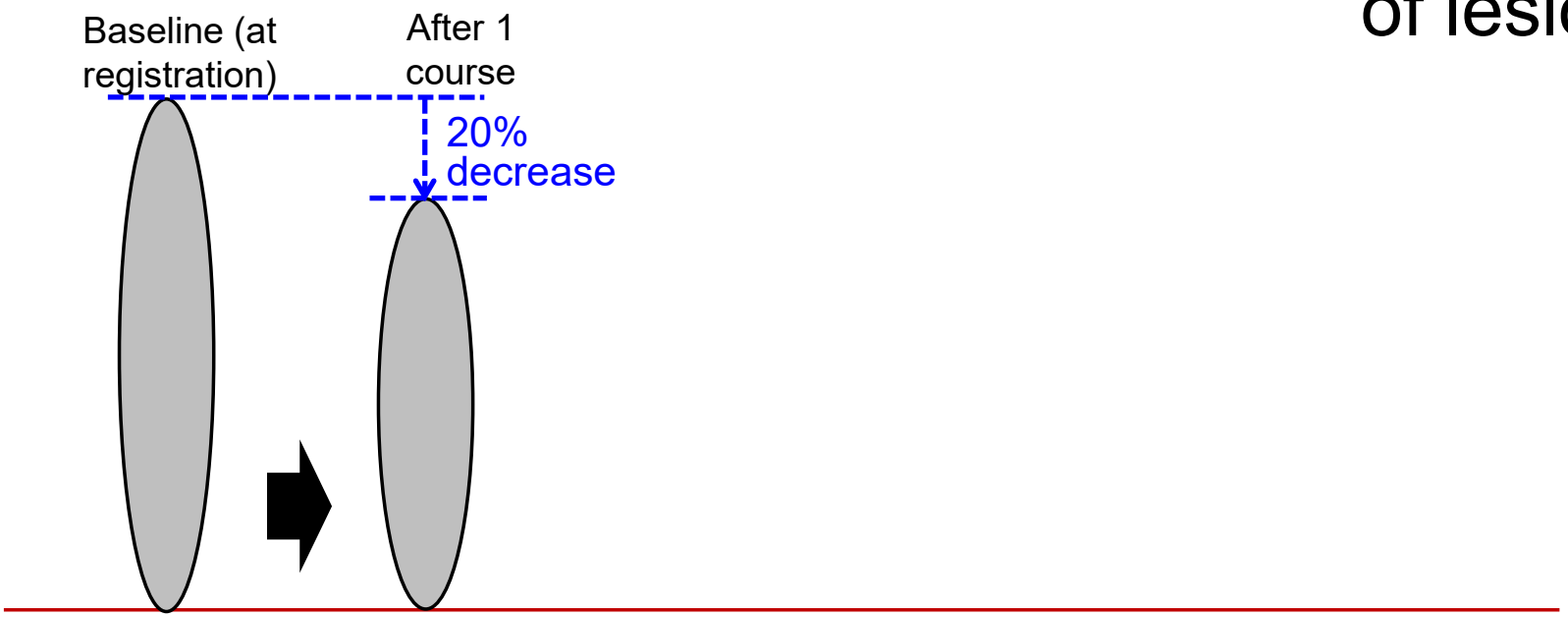
Response evaluation of target lesion:

supplement on PR and PD



Hypothetical example of response evaluation of lesion

TRY

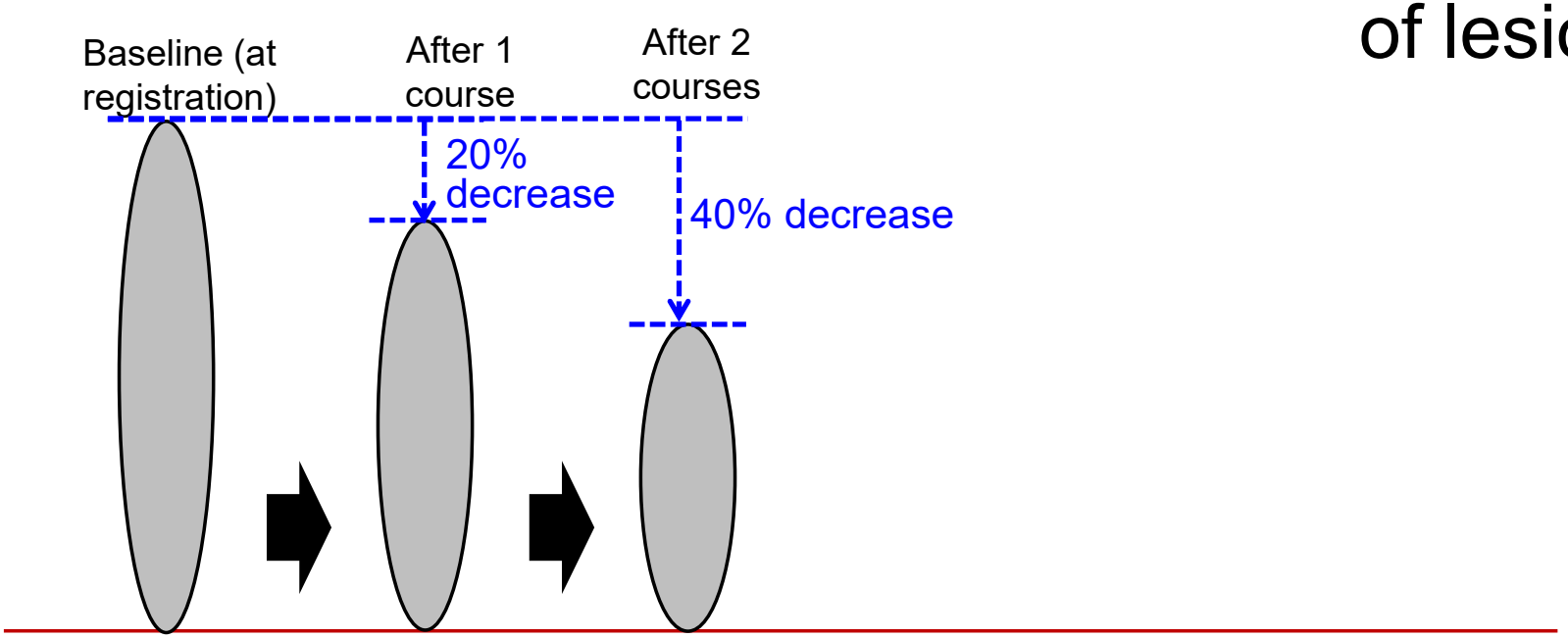


Sum of diameters	10 cm	8 cm
Decrease / increase rate in sum of diameters		(10−8)÷10 ＝20% decrease
Target lesion effect		SD

Decrease rate ＝ Sum of diameters at baseline is the denominator (PR for a decrease of 30% or more)
Increase rate ＝ Minimum sum of diameters is the denominator (PD for an increase of 20% or more)

Hypothetical example of response evaluation of lesion

TRY



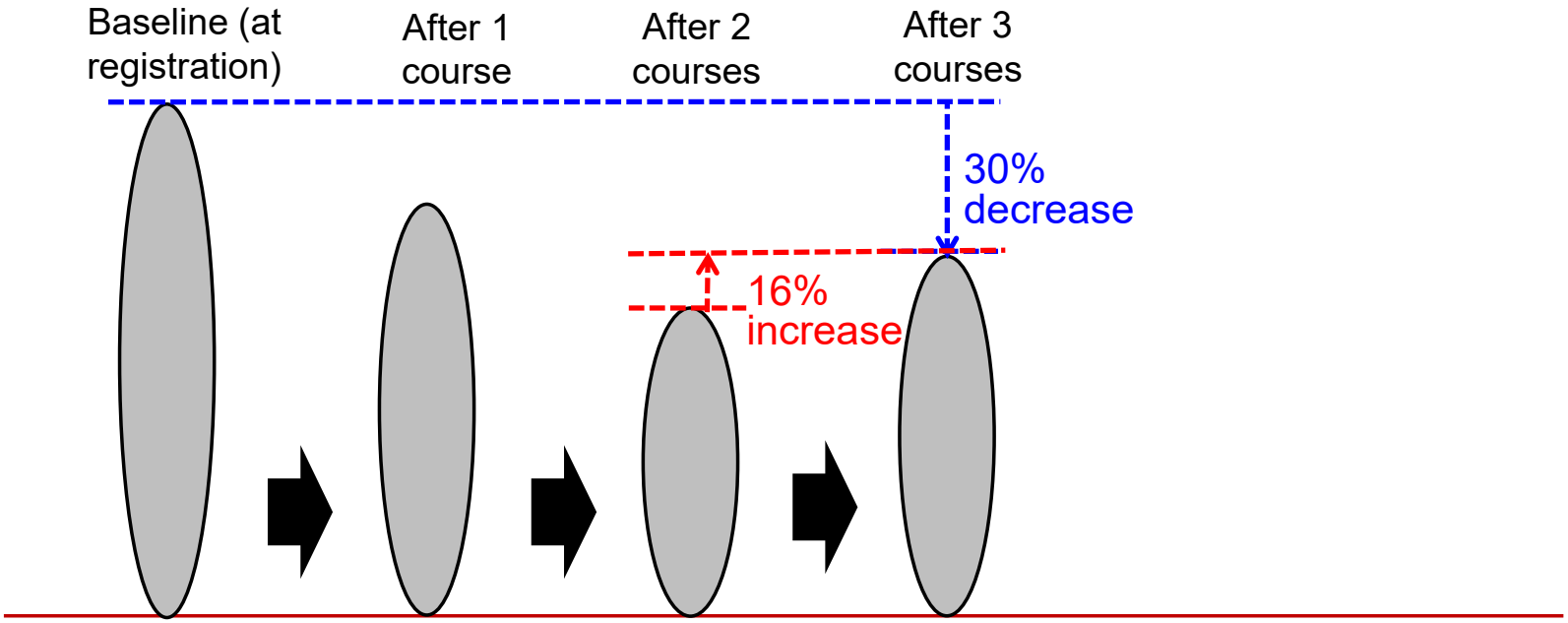
Sum of diameters	8 cm	6 cm
10 cm		
Decrease / increase rate in sum of diameters	(10−8)÷10 =20% decrease	(10−6)÷10 =40% decrease
Target lesion effect	SD	PR

Decrease rate ＝ Sum of diameters at baseline is the denominator (PR for a decrease of 30% or more)
Increase rate ＝ Minimum sum of diameters is the denominator (PD for an increase of 20% or more)

Hypothetical example of response evaluation

National Cancer Center Japan

TRY



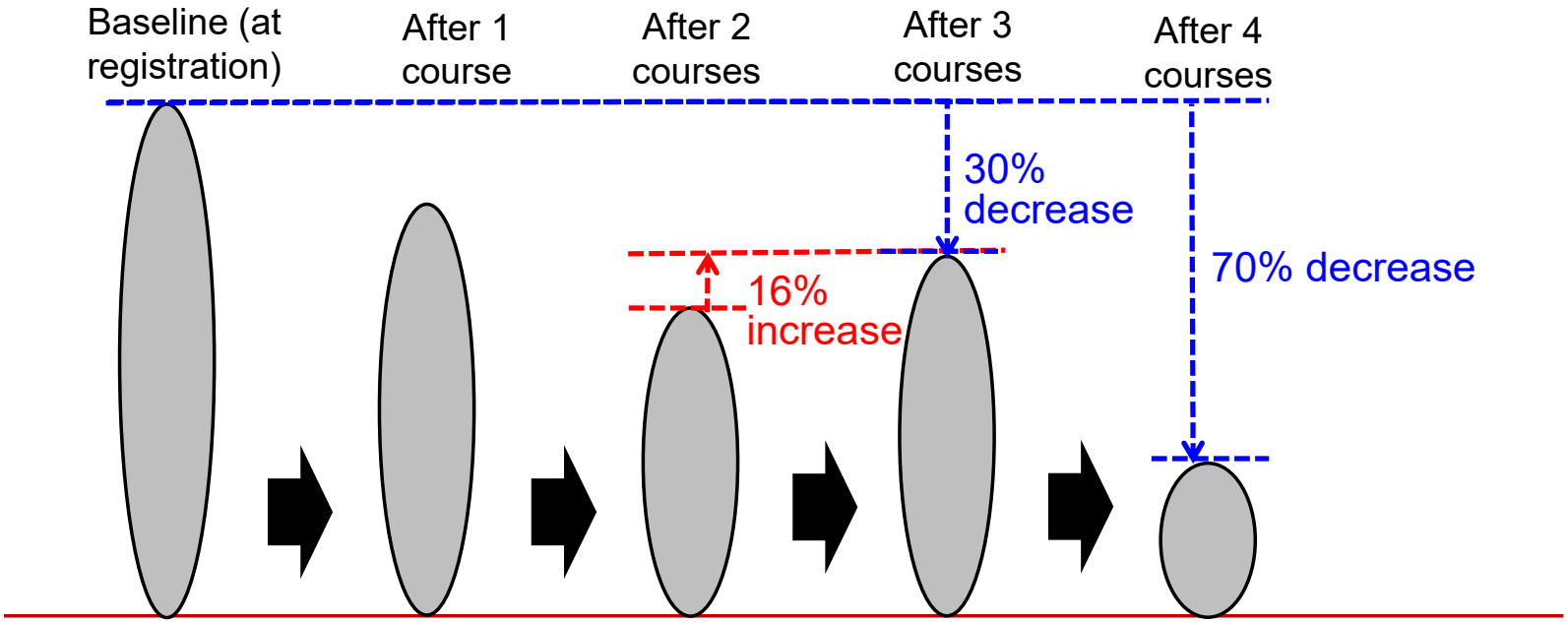
	Baseline (at registration)	After 1 course	After 2 courses	After 3 courses
Sum of diameters	10 cm	8 cm	6 cm	7 cm
Decrease / increase rate in sum of diameters		(10−8)÷10 ＝20% decrease	(10−6)÷10 ＝40% decrease	(10−7)÷10 ＝30% decrease (7−6)÷6 ＝16% increase
Target lesion effect		SD	PR	PR

Decrease rate ＝ Sum of diameters at baseline is the denominator (PR for a decrease of 30% or more)
Increase rate ＝ Minimum sum of diameters is the denominator (PD for an increase of 20% or more)

Hypothetical example of response evaluation

National Cancer Center Japan

TRY



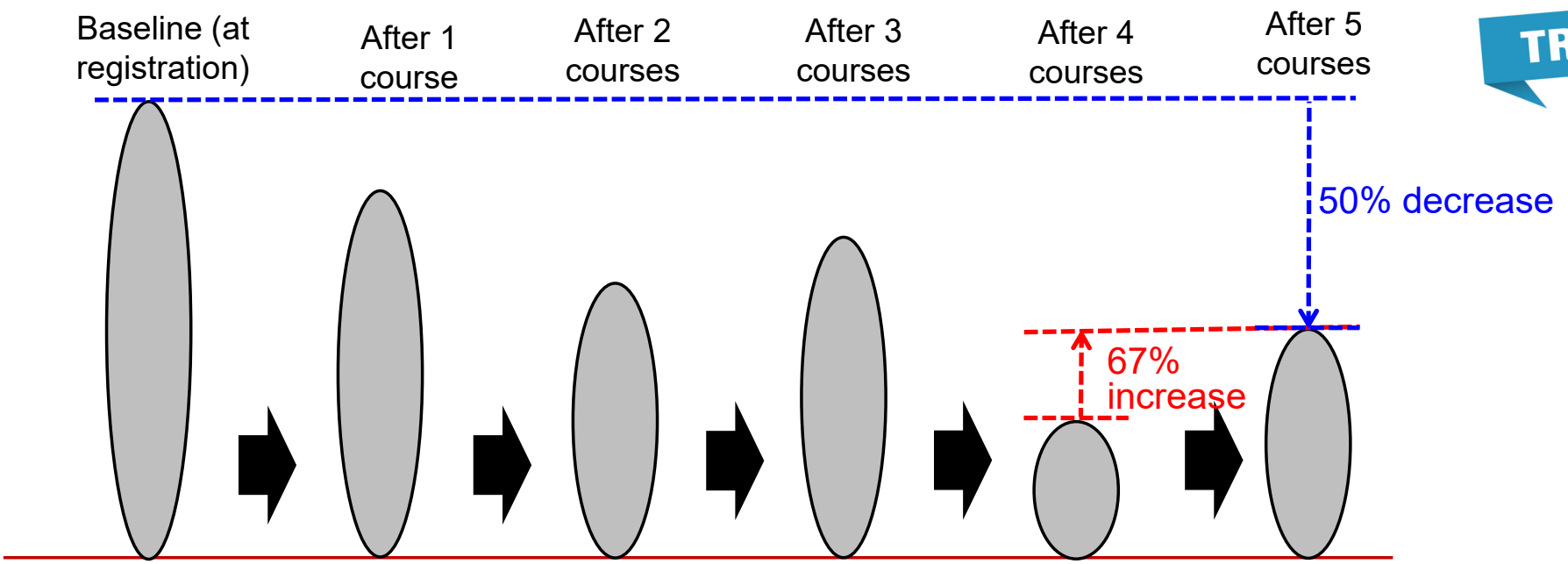
Sum of diameters 10 cm	8 cm	6 cm	7 cm	3 cm
Decrease / increase rate in sum of diameters	(10−8)÷10 =20% decrease	(10−6)÷10 =40% decrease	(10−7)÷10 =30% decrease (7−6)÷6 =16% increase	(10−3)÷10 =70% decrease
Target lesion effect	SD	PR	PR	PR

Decrease rate ＝ Sum of diameters at baseline is the denominator (PR for a decrease of 30% or more)
Increase rate ＝ Minimum sum of diameters is the denominator (PD for an increase of 20% or more)

Hypothetical example of response evaluation

National Cancer Center Japan

TRY



Sum of diameters 10 cm	8 cm	6 cm	7 cm	3 cm	5 cm
Decrease / increase rate in sum of diameters	(10−8)÷10 =20% decrease	(10−6)÷10 =40% decrease	(10−7)÷10 =30% decrease (7−6)÷6 =16% increase	(10−3)÷10 =70% decrease	(10−5)÷10 =50% decrease (5−3)÷3 =67% increase
Target lesion effect	SD	PR	PR	PR	PD

Decrease rate ＝ Sum of diameters at baseline is the denominator (PR for a decrease of 30% or more)
Increase rate ＝ Minimum sum of diameters is the denominator (PD for an increase of 20% or more)

Note: if target lesions become very small

- PD for target lesions
  - Sum of diameters increased by 20% or more relative to the minimum value during the course

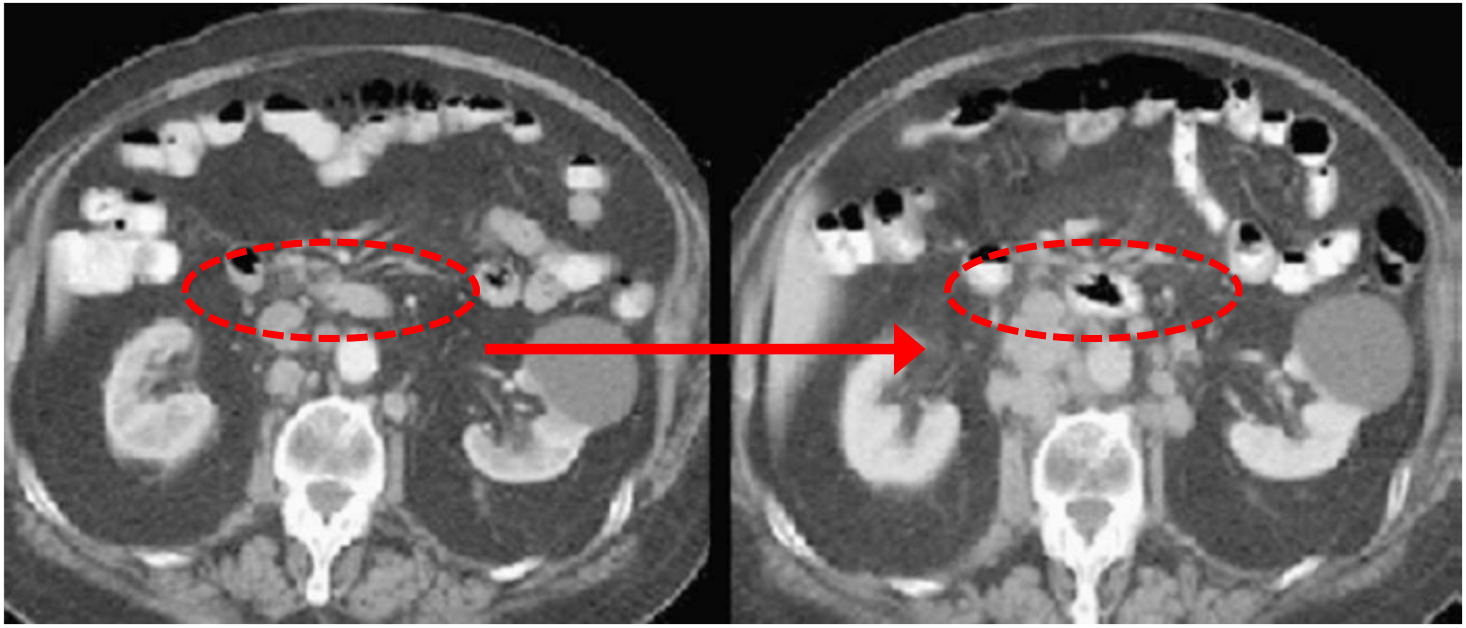
  and
  - Sum of diameters increased by 5 mm or more relative to the minimum value
    - e.g., sum of diameters is 20 mm → 24 mm ･･･is not PD
      (24－20)÷20＝20%　　24－20＝4 mm

- “Too small to measure”
  - Diameter of 5 mm or less
  - Record actual measurements as much as possible
  - Non-measurable lesion present → Recorded as diameter of 5 mm
  - Lesion likely absent → Recorded as a diameter of 0 mm

Unequivocal progression

- PD of non-target lesions
  - “unequivocal progression”
    - Clear progression such that the increase in overall tumor growth is sufficient to warrant **treatment discontinuation**

      (the treating physician would feel the need to change therapy)
    - Does not imply an increase in the diameter of a single non-target lesion



Eisenhauer EA et al. EJC 2009;45:228-47.

New lesion presence

- Definition of new lesion
  - Must be unequivocal: clear occurrence of new lesions
    - May not be tumor if judgment changes depending on test type / method
    - Continue treatment and re-evaluate for suspected cases
  - FDG-PET
    - Baseline PET negative → PET positive: PD
    - Baseline PET → PET positive: ?
      - PD if confirmed by CT
      - Follow-up with CT for suspected lesions
      - Not PD if ruled out by CT

(4) Judgment of overall response

- Overall Response
  - Judgment for each determined course

Target lesion	Non-target lesion	New lesion	Overall response
CR	CR	No	CR
CR	Non-CR/Non-PD	No	PR
PR	Non-PD	No	PR
SD	Non-PD	No	SD
PD	Any	Any	PD
Any	PD	Any	PD
Any	Any	Yes	PD

(5) Judgment of the best overall response

Calculate one “best overall response” from all “overall responses”

TRY

- Confirmation of effect → Ensure that there is no measurement error
  - Confirmation required
    - Required for single-arm trials with response rate as their primary endpoint
    - For confirmation of effect:
      - CR should occur twice in a row (best overall response CR)
      - PR should occur twice in a row (best overall response PR)

	After 1 course	After 2 courses	After 3 courses	After 4 courses	After 5 courses
Overall response	SD	PR	PR	CR	PD

→ best overall response **PR**

  - No confirmation
    - Not required in randomized controlled trials or trials that include preoperative treatment
    - Best “overall response” through all courses = “best overall response”

	After 1 course	After 2 courses	After 3 courses	After 4 courses	After 5 courses
Overall response	SD	PR	PR	CR	PD

→ best overall response **CR**

Summary of RECIST

- Purpose of RECIST
  - To “precisely” compare response rates in single-arm Phase II trials with historical controls
    - Also used to evaluate progression-free survival (PFS) in addition to response rate
  - Distinguish between PR and SD with a common global standard, in order to improve comparability
  - Not an index for deciding whether to continue / discontinue treatment
    - Consider response evaluation and individual patient treatment continuation / discontinuation separately
  - RECIST is a “guideline” rather than response evaluation criteria themselves
    - Imaging modalities used, timing of judgments, definition of measurable lesions, presence / absence of confirmation, and other details must be specified in the protocol for each trial
  - There are different response evaluation criteria classifications, but the basic idea is the same
    - RANO, Lugano, etc.