PRO estimands in oncology: from treatment to beyond *Some methodological notes*

Can the Estimand Framework Strengthen Patient-Centered Cancer Research Throughout the Evidence Generation Lifecycle Beyond Approval: Opportunities and Challenges Ahead?

Konstantina Skaltsa, *Senior Director, Research and Scientific Affairs, Patient-Centered Solutions*

May 19, 2026

PRO objectives in oncology: Why estimands matter

Have we made the right effort to define the objectives precisely enough?

Clinical objectives are thoroughly discussed across multidisciplinary teams when designing a study
→ Are we doing the same with PROs?
Objective wording such as "to assess HRQoL, functioning and symptoms in patients with … " is still a common wording in protocols.

Have we carefully evaluated the right schedule of assessment for PROs?

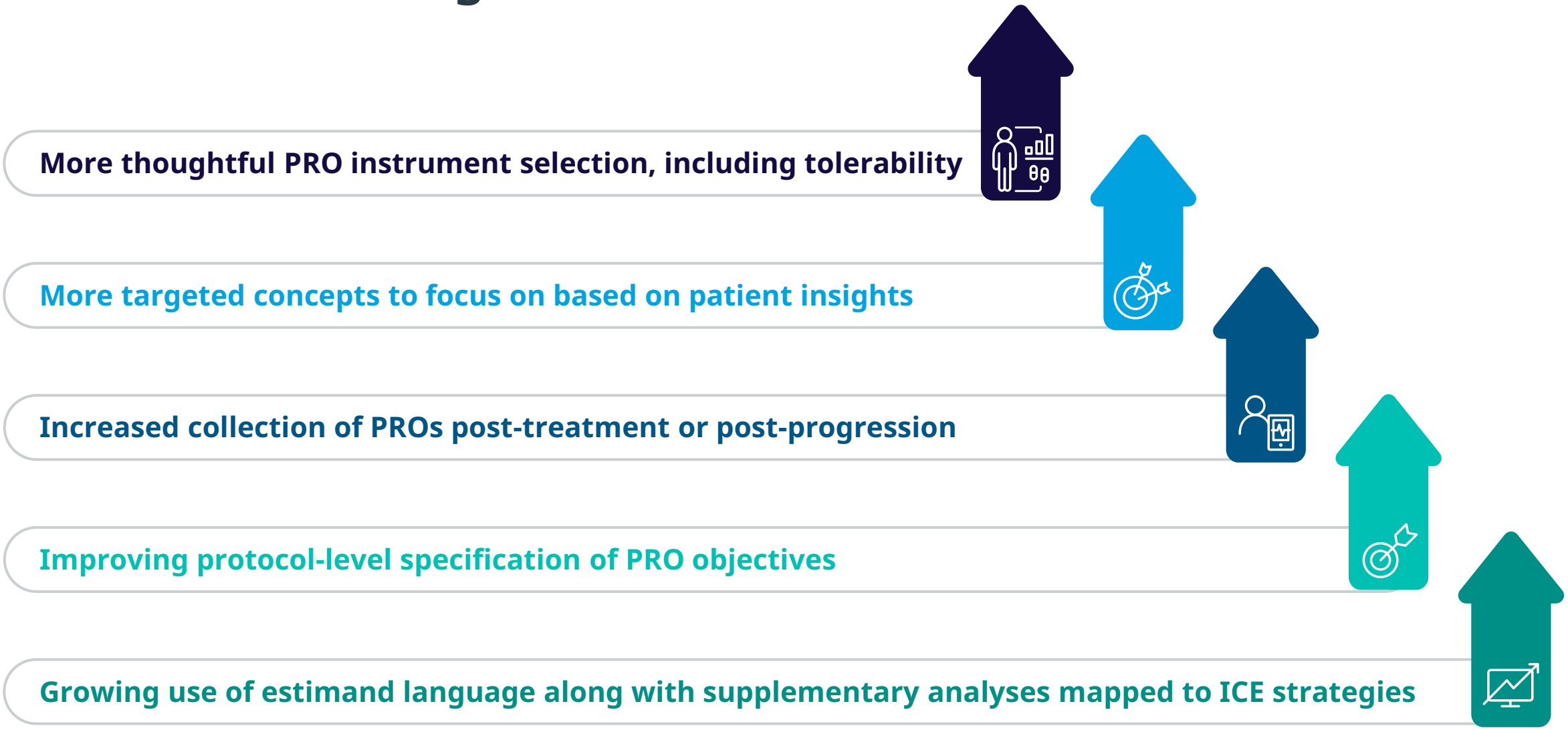
Common to consider convenient collection at the beginning of each cycle and every X weeks after treatment discontinuation for EQ-5D only with a budget-limit of X assessments ("payers are requesting it")

Are we using the estimand framework to define our endpoints precisely?

The estimand framework offers a structure to align:
- **Scientific questions**
- **Endpoint definition**
- **Analysis strategies**

HRQoL: Health Related Quality of Life; PRO: Patient Reported Outcome

What are we doing well?



ICE: Intercurrent Event; PRO: Patient Reported Outcome

There are some generic issues that still hinder the definition of well-crafted objectives

1 Considerable divergence between stakeholders reviewing the same clinical trial data

- Instruments used
- Time horizons
- Concepts prioritization
- Common practices include:
  - **Over-collecting**, *e.g., concept overlap between instruments, aim for collecting all PROs for as long as possible etc.*
  - **Under-collecting**, *e.g., while an endpoint targeting a specific timepoint has been defined, PRO collection ends after treatment discontinuation (which may occur before this timepoint)*

2 eCOA implementation and compliance issues are still common

3 Lack of insights on reasons for missing and plausible PRO values after ICEs

Despite massive calls for years, instead of looking for evidence to equip our statisticians with well-informed assumptions, we resort to familiar statistical methods *"that do not rely on assumptions"...*

eCOA: electronic Clinical Outcome Assessment; ICE: Intercurrent Event; PRO: Patient Reported Outcome

And there are specific methodological issues we are still not ready to accept solutions for

1. **MMRM with time considered a categorical variable may not be appropriate in complex SoA**

2. **Flat MMRM for all concepts treats all scores/concepts in the same way *("ensures consistency")***

3. **Flat MMRM treats missing data after all ICEs in the same way: hypothetical strategy**

4. **Treatment policy is desired by everyone but implemented by no one**

SoA can be different between the two arms

- Depending on cycle length
- Less frequent after treatment discontinuation, which can occur at different timepoints per patient

Could it be that when stopping treatment

- Dyspnea gradually and continuously worsens?
- Nausea goes down, e.g., back to baseline levels or even resolve entirely?
- Physical functioning remains stable for a certain period, e.g., a year?

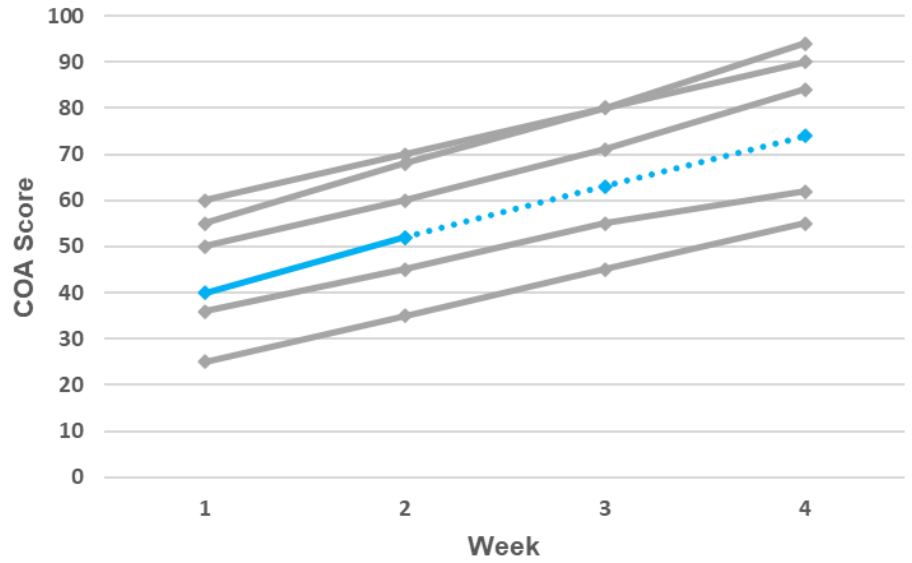
Unclear objective post-treatment discontinuation

- How often and until when to collect PROs?
- Which PROs?
- Often thinking about the estimand and the data it needs occurs mid-trial --> amendments possible, but difficult and reduce sample

ICE: Intercurrent Event; MMRM: Mixed Models for Repeated Measures; PRO: Patient Reported Outcome; SoA: Schedule of Assessments

Historic use of MMRM

- Introduction of the repeated measures framework constituted an improvement over LOCF or AUC approaches reigning before it
  - Promise was that “all the repeated measures can be used and missing data handled”
  - This treats missing data as one unique issue
- Took some time to understand the assumption of “similarity”
  - When we did, we started discussing sensitivity analyses for the possibility of MNAR
- Estimands pushed us into a different angle: not all missing data are occurring due to the same reason
  - Pattern-mixture models have pushed the thinking into splitting the analyses “by reason”



Under MAR, the MMRM model estimates the mean treatment effect assuming that “*. . . after withdrawal, subjects* ***would have continued*** *just like their peers in the same arm who have the same covariates and same observed data (so far)”.*

We are not yet carefully considering each reason for unobserved data or each concept separately

Quote by James Roger. https://www.psiweb.org/docs/default-source/resources/psi-subgroups/scientific/2015/estimands-28-09-2015/jamesroger.pdf?sfvrsn=bba3d0db_2

AUC: Area Under the Curve; LOCF: Last Observation Carried Forward; MAR: Missing At Random; MMRM: Mixed Model for Repeated Measures; MNAR: Missing Not At Random

All estimand strategies are assumption-heavy, but some cause more dysphoria than others

Hypothetical strategy

- Traditionally (and unintentionally) used
- Assumptions implausible, however, most frequently used
- Strategically, but not scientifically, may be a reasonable starting point

While-on-treatment strategy

- Term often misused interchangeably to the hypothetical strategy as *"only while-on-treatment data are used in the model"*
- Appropriate estimators are not clear in the field

Composite strategy

- May align more closely to traditional regulatory preference for "conservative" estimates of treatment effect
- Challenging to implement when analyzing continuous variables

Principal stratum strategy

- Originating from the causal inference framework: promising but largely unfamiliar in the field
- Data-sensitive and assumptions considered harder to swallow than other strategies

Treatment policy strategy

1. Aims to reflect real-world experience more closely
2. Includes PROs after ICEs
3. Sensitive to missing data and follow-up burden

Treatment policy is easy to say and seems to be enjoying consensus, but requires clarification

For what concepts?

- Efficacy objectives: generally, yes. Includes:
  - HRQoL
  - Functioning
  - Disease-related symptoms
- Tolerability objectives – perhaps while-on-treatment?
  - Treatment-related

What about those that are not clearly disease- or treatment-related?

For what ICEs?

- Treatment discontinuation:
  - Yes, if efficacy objective
- Treatment switching?
  - Wishes unclear
  - Collection feasibility considerably harder
- Death → not possible

For how long?

- The FDA would look for specific timepoints with thorough justification of selection
- The EMA would look for longer horizons
- The HTAs would look for even longer horizons, where data are scarce

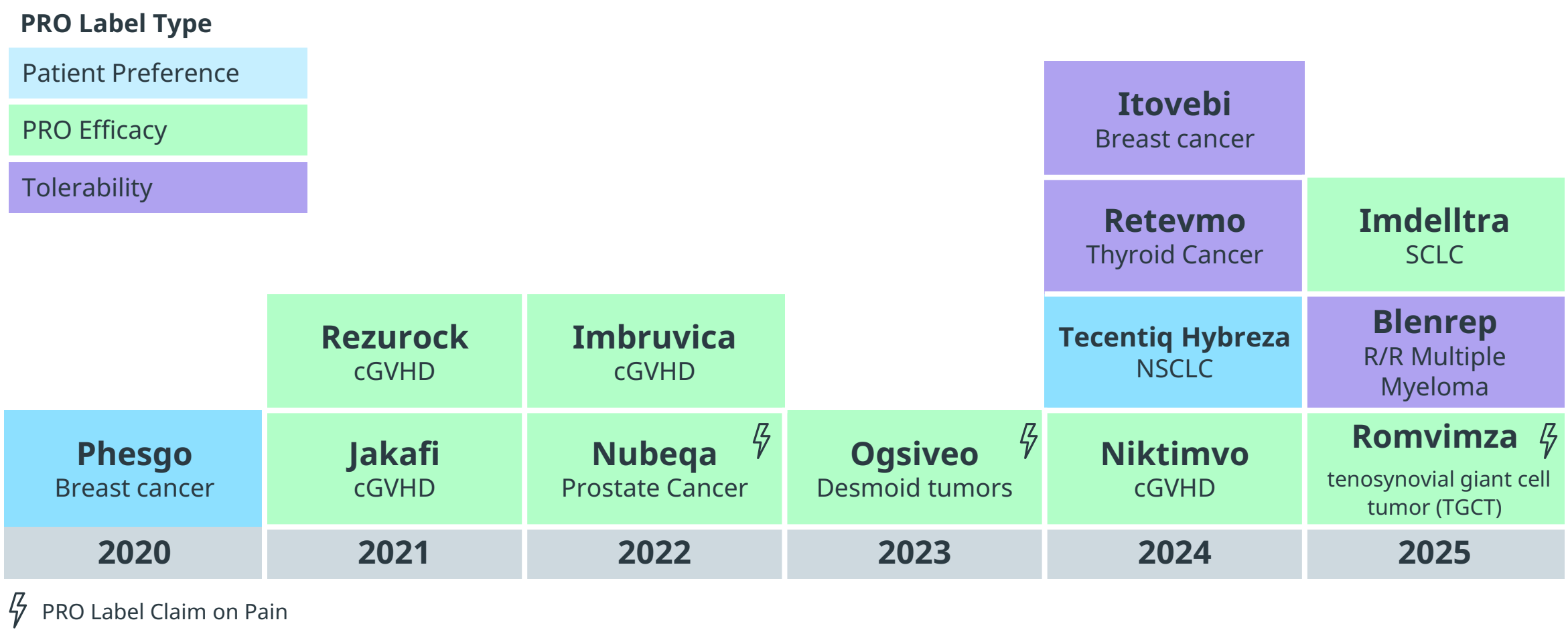
Important SoA/design implications

Methodological complexity when data are scarce

EMA: European Medicines Agency; FDA: Food and Drug Administration; HRQoL: Health-Related Quality of Life; HTA: Health Technology Assessment; SoA: Schedule of Assessments

On a positive note, since 2020, several PRO labels have been achieved in oncology products

Over the past 2 years, there has been an increased attention by the FDA to examine and provide labels for patient-reported tolerability in addition to PRO efficacy (e.g., disease-related symptoms)



FDA Project Patient Voice

Key points

1 **Define PRO objectives upfront**
Using the estimand framework, including post-treatment questions, rather than retrofitting estimands mid-trial

2 **Plan PRO collection deliberately**
Prioritizing concepts, time horizons, and ICEs that truly matter to patients, regulators, and HTA bodies - avoiding both over- and under-collection

3 **Move beyond one-size-fits-all analyses**
Selecting estimand strategies and models that reflect the clinical reality of different PRO concepts and intercurrent events

4 **Be explicit about assumptions**
Especially around missing post-ICE data; transparency is preferable to relying silently on implausible defaults

5 **Translate consensus into implementation**
Recognizing that treatment policy estimands require investment in design, follow-up, and methodological rigor to be credible and usable.

Thank you!

If you have questions or comments, please do reach out to:
Konstantina.Skaltsa@iqvia.com