

The next frontier in patient reported outcome assessment: AI-driven and chatbot-based measurement and evaluation

Professor Mogens Groenvold, MD, PhD

Copenhagen University Hospital, Bispebjerg/Frederiksberg, Denmark

OEI Oncology Days, Warsaw, 17 June 2026

Agenda

- **How can AI, GAI and chatbots be used** in the assessment of PROMs?
 - Should we expect AI, GAI and chatbots to **replace the current approaches** to the assessment of PROMs – such as the EUonQoL Tool-Kit?
 - What are the **clinical implications and future directions** for PROMs research?

Current state of PROMs/AI

- AI used extensively in healthcare – mainly for data already available (omics, imaging data, medico-administrative data) but very little for PROMs
- PROMs are essential but often fragmented, inconsistently collected, and under-used in research and clinical practice
- The lack of attention to PROMs/AI:
 - May further marginalise PROMs
 - Impairs further research
 - Marginalises those are not represented



How can AI, GAI and chatbots be used in the assessment of PROMs?

Google AI (12 June 2026): ‘AI, Generative AI (GAI), and chatbots are transforming Patient-Reported Outcome Measures (PROMs) by replacing static, rigid questionnaires with dynamic, personalized conversational interfaces. These technologies enhance accessibility, reduce patient burden, and provide deeper, real-time insights into a patient's health status and quality of life. [[1](#), [2](#)]’

1. Boyer et al. NPJ Digital Medicine 2025: 8; 624.
2. Terheyden et al. Journal of Patient-Reported Outcomes (2025) 9:34

‘LLM-PROM’ (Terheyden)

- ‘A large language model is prompted based on an existing, validated patient-reported outcome measure (source PROM) and additional text bodies.’
- ‘A machine learning algorithm is trained to estimate the LLM-PROM score based on individuals’ free-text responses based on a user interaction and source PROM scores.’
- Could
 - Improve adoption
 - Increase comprehensibility
 - Increase content validity
- **Could create a ‘black box’ process**
- **Requires extensive validation**



A typical PROM item

During the past week:

Have you had pain?

- Not at all
- A little
- Quite a bit
- Very Much

Imaginary conversation *(prompted in Google AI, first attempt)*

Doctor: Hello. I see from your chart that you have been experiencing significant pain. Let us focus on how you have been feeling over the past week. How exactly did this pain begin seven days ago?

Patient: It started last Monday as a dull, constant ache. It was manageable at first, but it just has not gone away.

Doctor: Has this pain remained in the exact same spot all week, or has it started to radiate or move to other parts of your body?

Patient: It actually started in my lower back, but over the last few days, it has started shooting down into my right leg.

Doctor: Looking back over the last seven days, has the pain been constant, or does it come and go in waves?

Patient: The ache in my back is always there, but the sharp shooting pain down my leg comes and goes, especially when I try to move around.

Doctor: On a scale of 1 to 10, with 10 being the worst pain imaginable, how would you rate the pain at the start of the week versus right now?

Patient: It started as a 3 or 4 early in the week. Right now, as I sit here, it is easily an 8.

Doctor: Have you noticed any specific activities or positions over the past week that make the pain better or worse?

Patient: Sitting down or lying flat makes it much worse. Walking around slowly seems to provide a little bit of relief.

Doctor: Have you taken any pain relief medication during the week, and if so, how much did it help?

Patient: I have been taking ibuprofen every day, but it only takes the edge off for an hour or two before the pain returns completely.

Imaginary conversation (*prompted in Google AI*)

Doctor: Hello. I see from your chart that you have been experiencing significant pain. Let us focus on how you have been feeling over the past week. How exactly did this pain begin seven days ago?

Patient: It started last Monday as a dull, constant ache. It was manageable at first, but it just has not gone away.

Doctor: Has this pain remained in the exact same spot all week, or has it started to radiate or move to other parts of your body?

Patient: It actually started in my lower back, but the last few days, it has started shooting down into my right leg.

Doctor: Looking back over the last seven days, is the pain constant or does it come and go in waves?

Patient: The ache in my back is always there, but the leg pain comes and goes, especially when I try to move around.

Doctor: On a scale of 1 to 10, with 1 being the least pain at the start of the week versus right now?

Patient: It started as a 3 or 4 early in the week, but now it's a 7 or 8.

Doctor: Have you noticed any specific activities or positions that make the pain better or worse?

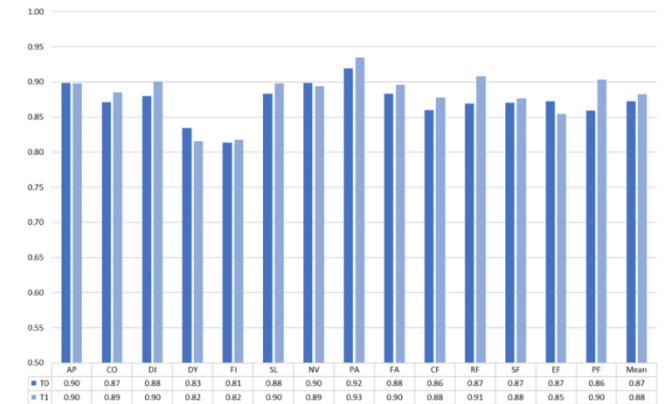
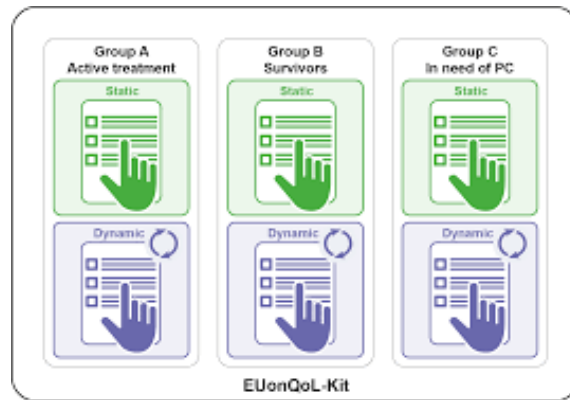
Patient: Sitting down or lying flat makes it much worse. Walking around gives me a little bit of relief.

Doctor: Have you taken any pain relief medication during the week, and if so, how much and how often?

Patient: I have been taking ibuprofen every day, but it only takes the edge off for an hour or two before the pain returns completely.

**NB. Second version.
Had to correct the chatbot, which
suddenly started to ask about
abdominal pain rather than pain**

How could this happen? Current practice

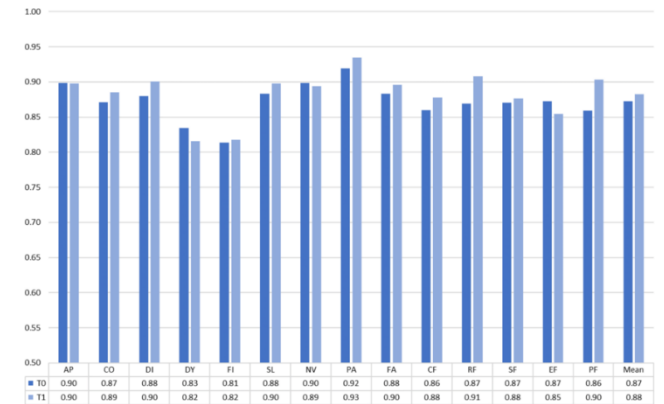
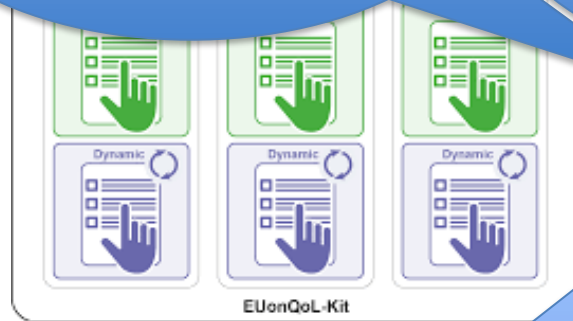


How could this happen? Future practice, within system

AI increases
completeness
of data

AI increases
accessibility for
vulnerable pts.

AI increases
validity, eg. by
exploring
inconsistencies

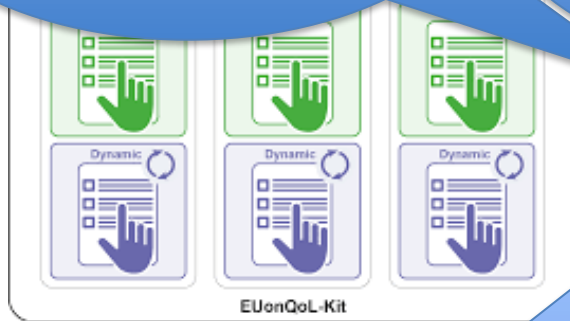


How could this happen? Future practice within system

AI increases **completeness** of data

AI increases accessibility for vulnerable pts.

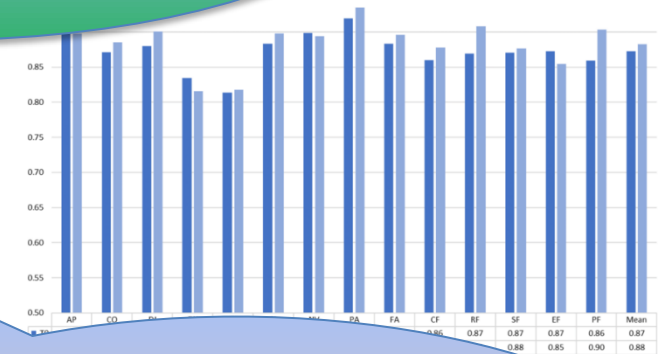
AI chatbot adds relevant questions based on responses and 'insight' in other data



AI increases **validity**, eg. by exploring inconsistencies

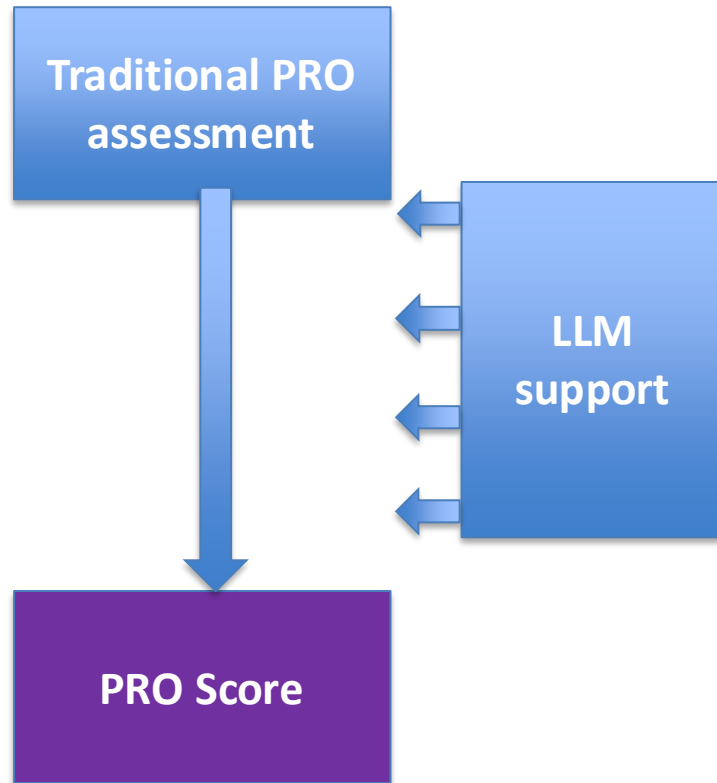
AI chatbot adds rich qualitative data

AI and GAI assists HCP to provide more pt-centered care based on analyses combining enhanced PRO data and other data



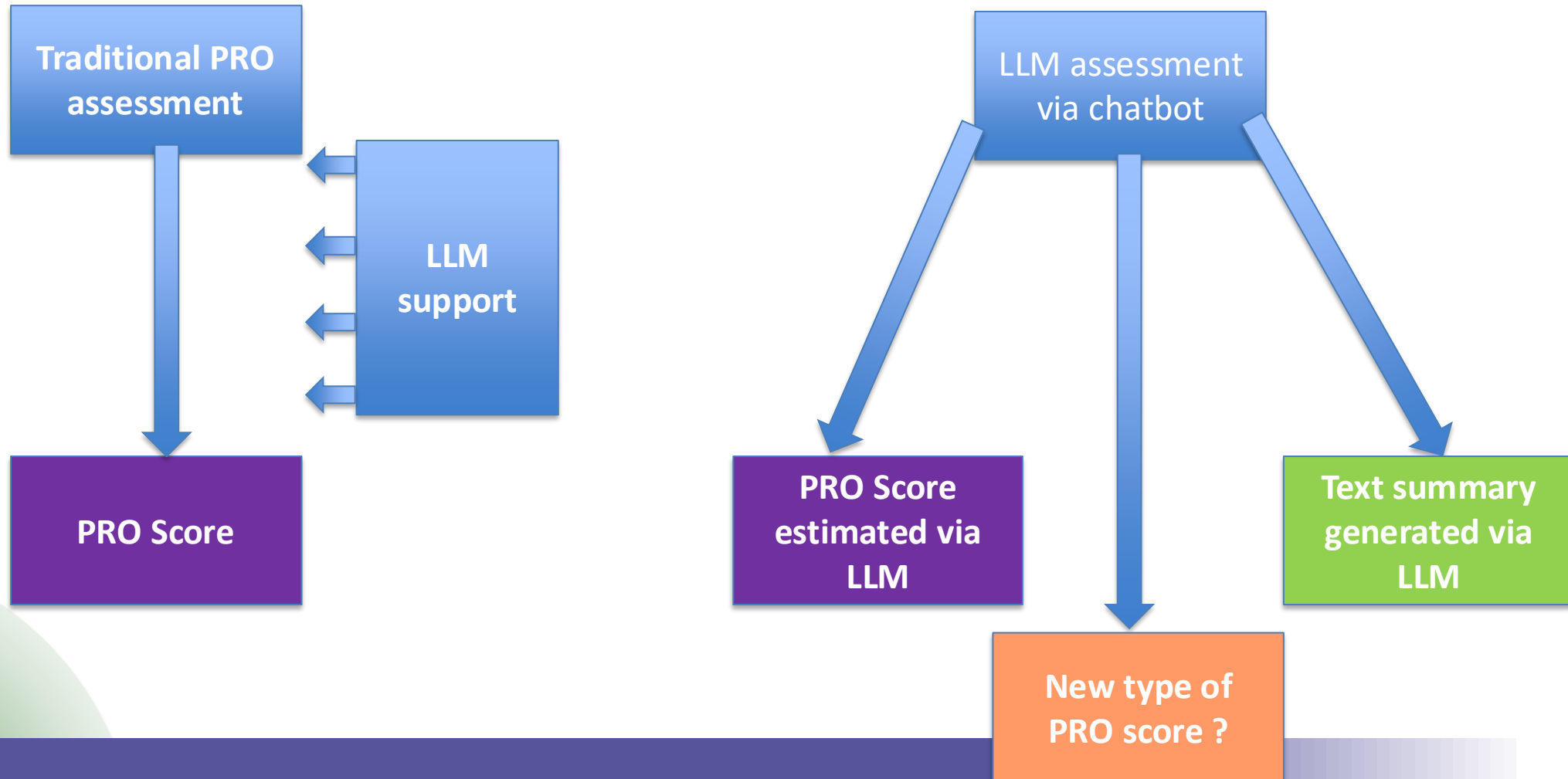
Summary focusing on the **assessment**

(similar models can be made for selection of domains, analytical integration with other data, etc.)



Summary focusing on the **assessment**

(similar models can be made for selection of domains, analytical integration with other data, etc.)

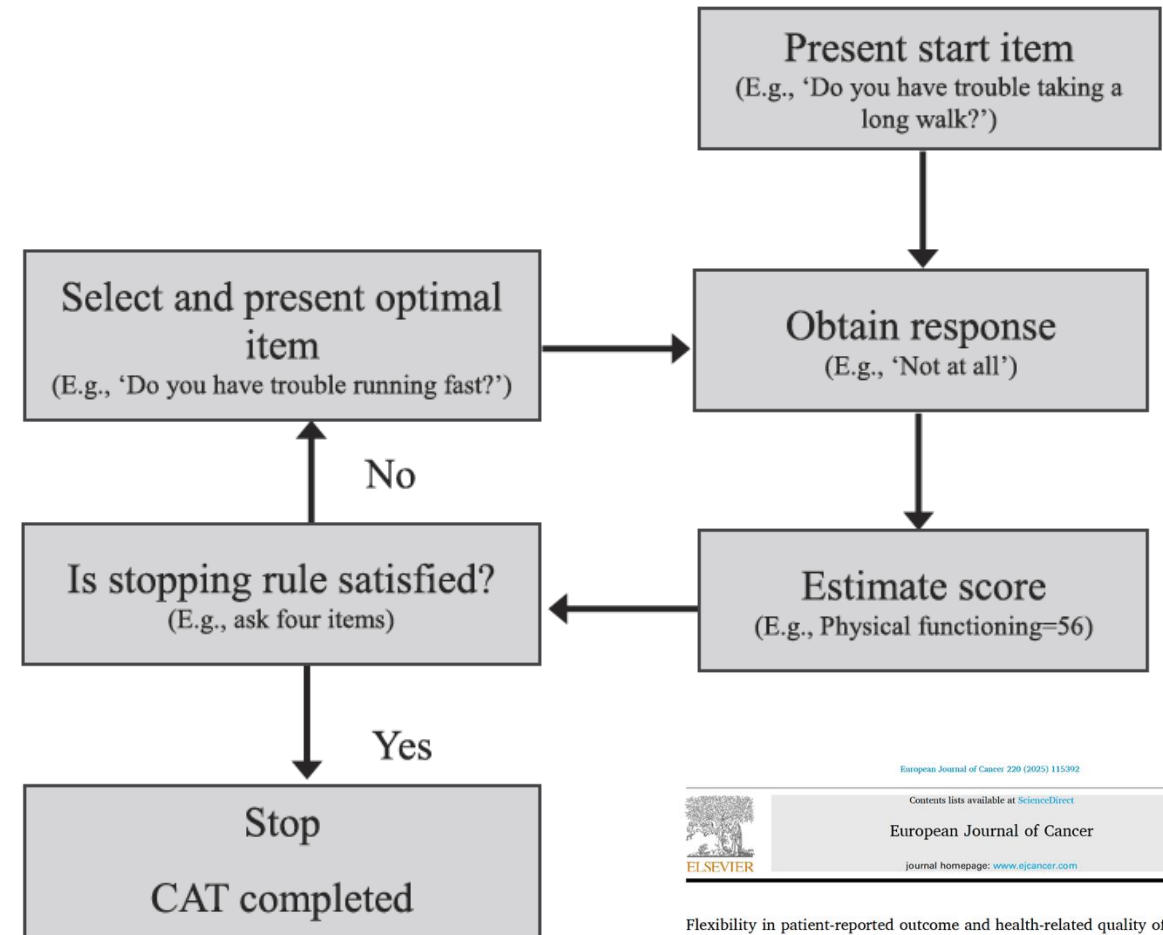


We need to consider Psychometrics

- Psychometrics is the methodological basis for PROMs:
- **Classical test theory**
 - Assumes all items have equal weight
 - A scale score is the sum (mean) of responses
- **Item Response Theory (IRT) and Rasch models**
 - Does not assume that all items 'are equal'; rather the 'difficulty of each item is estimated (e.g., can walk 100 m vs. being able to run a marathon)
 - Allows for **Computer-adaptive Testing (CAT)**

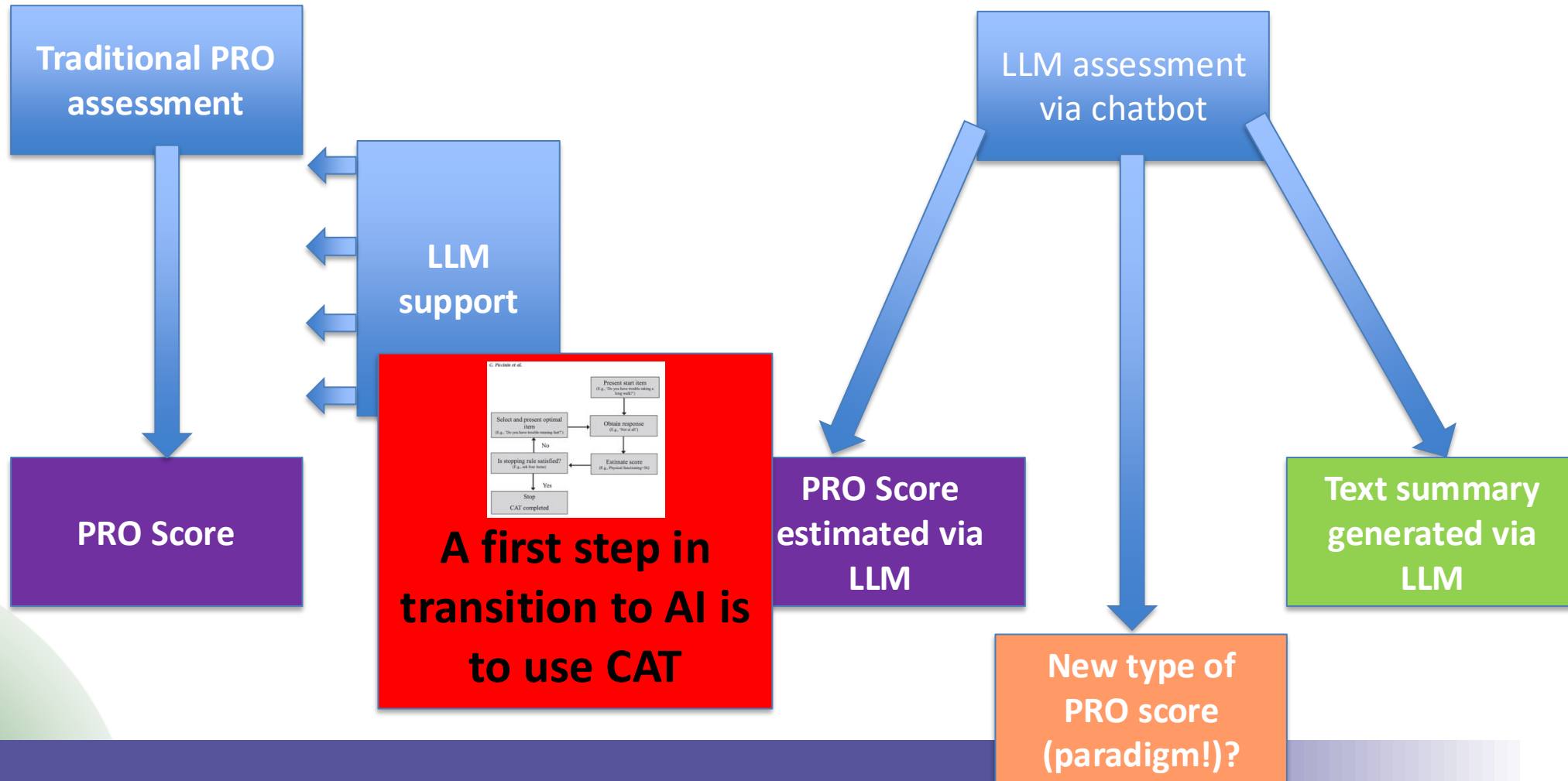
Computer-adaptive testing as the logical first step

- For each domain, such as physical functioning, an 'item bank' is developed
- Each item is characterized by the 'difficulty' of the task it asks about, e.g. length of walk
- Selection of items is individualised
- Scores are represented on broader and more detailed scale of measurement
- Higher precision, less floor/ceiling effect, lower sample size requirement



Summary focusing on the **assessment**

(similar models can be made for selection of domains, analytical integration with other data, etc.)

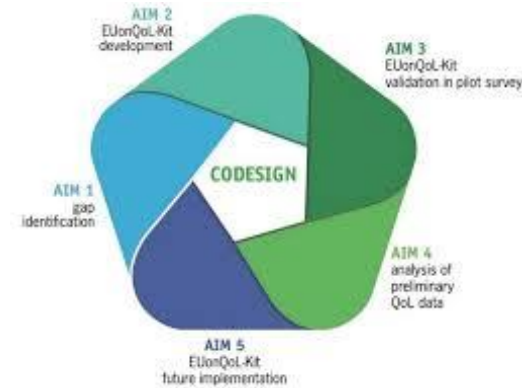


Some of the clinical perspectives

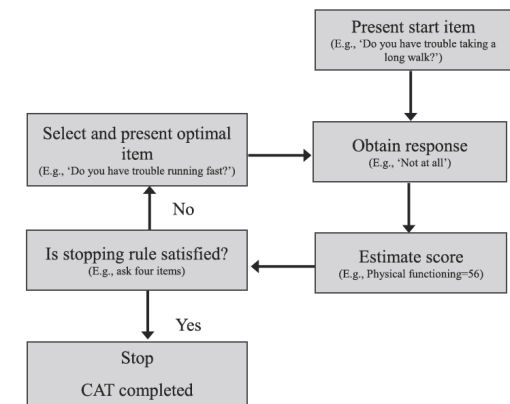
- ‘Making them (PROMs) **more accurate**, through true personalization that better reflects each patient’s unique experience;
- **more meaningful**, by enabling qualitative insights that capture the richness of lived health narratives; and
- **more actionable**, by producing clinically relevant information that can be seamlessly integrated into care.’
 - Boyer et al. NPJ Digital Medicine 2025: 8; 624.

So, what does this mean for EUonQoL?

- The Toolkit is a contemporary PROM based on co-creation
- Some parts of Toolkit (8 domains such as physical and emotional function, fatigue) available for computer-adaptive Testing (CAT)
- Many of the proposed AI-based innovations may increase the user experience related to Toolkit, simplify the response process, lead to fewer errors and missing data, and make it more accessible to vulnerable groups of patients



C. Piccinin et al.



Conclusions

- AI, GAI and chatbots may to improve the quality, completeness and relevance of PRO data in all stages of the research process
- Need for extensive research for this to happen – we cannot take any of the ‘promised’ improvements for granted
- Must build on the rigorous methodology already available (psychometrics, validation, etc.), which needs to evolve and adapt
- Quick and decisive action is needed: PROs have been marginalised for too long – we may miss another train if PRO data do not get a larger role in the ongoing AI-based developments in health care

Thank you!



mold@sund.ku.dk

Mogens Groenvold, MD, PhD

Copenhagen University Hospital, Bispebjerg/Frederiksberg, Denmark

Definitions

- ‘**AI**, or artificial intelligence, refers to computer systems designed to perform tasks that typically require human intelligence, such as recognizing patterns, making decisions, understanding language, and solving problems.’
- ‘**Generative AI (GAI)** is a subset of artificial intelligence that creates new content such as text, images, music, code, or video based on patterns from the existing data. Unlike traditional AI, which primarily analyzes data and makes predictions or classifications, Generative AI actually produces original outputs. ’
 - <https://www.transorg.ai/blog/a-guide-to-understand-ai-gai-ml-llm-gans-and-gpts/>
- A **chatbot** is a software application designed to simulate human conversation