From Lived Experience to Longitudinal Evidence: Designing an OMOP-Compatible Export Layer for Prospective Cancer Patient-Reported Outcomes in Sweden

Máté Bendegúz Horváth(hrvtmate10@gmail.com)[1], Carlos Manriquez(carlos@digitalpotential.se)[1], Lars Halvorsen(lars.halvorsen@edence.health)[2], Christian Högberg(christian@passion2improve.se)[3]
[1]Oncoly AB, Sweden
[2]edenceHealth, Kontich, Belgium
[3]Passion 2 Improve Sverige AB, Sweden

Background

When a cancer patient leaves the clinic after chemotherapy, they enter a space largely invisible to research. Symptoms are common among individuals with cancer, yet many go undetected by clinicians, even in controlled clinical trials, physician under-reporting of patient-experienced toxicity ranges from 41% to 74% [1,2]. This gap exists not because clinicians fail to ask, but because the healthcare system was never designed to listen continuously.

The OHDSI community has recognized this limitation explicitly. In 2022, van Sandijk and colleagues observed: "There is high demand for high-quality longitudinal datasets with information about treatment effects on patients' daily life. Currently, however, OMOP-CDM and OHDSI tools do not really support PROMS data" [3]. Three years later, this gap persists. A 2025 scoping review of OMOP adoption in cancer research found extensive work on clinical registries and EHR transformation, but prospective patient-reported outcome platforms remain notably absent from the standardization landscape[4].

The invisibility extends beyond patients. Informal caregivers like spouses, children, parents who become round-the-clock support systems provide care that is vital yet systematically uncaptured. Research demonstrates that caregiver burden directly correlates with patient outcomes, yet validated instruments like the Zarit Burden Interview rarely appear in observational databases[5]. Caregivers remain, as one meta-analysis described, "invisible" to healthcare systems despite shouldering immense responsibility[6].

A third gap compounds the problem: hereditary cancer risk. When family history is incomplete, which occurs in the majority of cases patients who should receive genetic testing referrals are missed. Del Fiol and colleagues found that augmenting EHR data with comprehensive family history more than doubled identification of individuals meeting genetic evaluation criteria, and quadrupled detection among those with extensive family records[7]. Yet prospective, systematic family history collection is rarely integrated with symptom monitoring.

Oncoly was founded by a brain-cancer survivor who experienced these gaps. The platform captures what happens between clinic visits: weekly symptom logs, EORTC QLQ-C30 quality-of-life assessments, ZBI-6 caregiver burden measures, and family cancer history, all prospectively, with explicit consent for research use. The project, aiming to secure funding through Sweden's MedTech4Health program and aligned with the OMOP 4 Sweden! initiative, aims to design an OMOP-compatible export layer that transforms this uniquely comprehensive dataset into a standardized resource for the research community. OMOP 4 Sweden! is an initiative led by Passion 2 Improve Sverige AB and financed by the Swedish Innovation Agency Vinnova through the Swelife[8] strategic innovation programme.

Methods

Oncoly creates value in two directions simultaneously. The platform delivers direct value to patients, caregivers, and clinicians, transforming collected PRO, quality-of-life, and hereditary indicators into actionable insights that support care between clinic visits. At the same time, standardizing this data to OMOP CDM creates indirect value by contributing prospective, consented patient experience data to the research community.

Our approach follows OHDSI's established methodology (1), planned to be implemented in collaboration with edenceHealth (Belgium) and aligned with the OMOP 4 Sweden! initiative supported by Swelife and Vinnova.. Figure 1 illustrates the data architecture: patient-reported symptoms, caregiver burden assessments, and family cancer history are all captured in Oncoly's production database, where they support clinical care.

This architecture ensures that research workloads never impact patient-facing operations while enabling continuous data contribution as new patients enroll. Future extension (2027) will add federated analytics capability, allowing privacy-preserving multi-site queries without centralizing patient-level data.

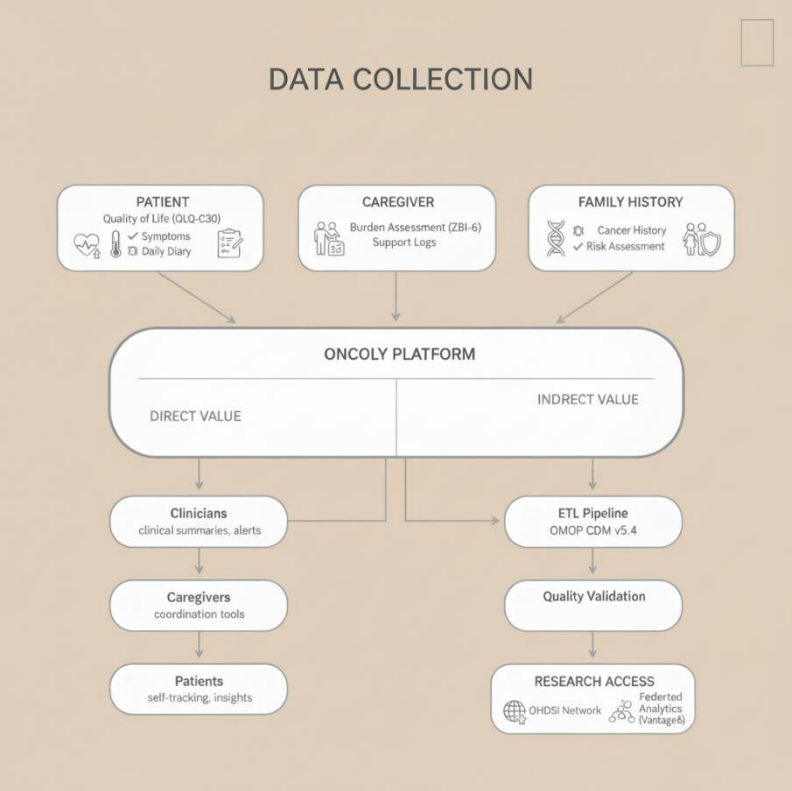

Figure 1. Oncoly dual-value data architecture showing direct clinical value creation (left pathway) and indirect research value via OMOP standardization (right pathway).

Results

With the planned conclusion of this project in late 2026, Oncoly will have established a replicable pathway for bringing prospective patient-reported data into the OHDSI ecosystem.

The first deliverable is an OMOP schema design that demonstrates how prospective PRO data, including the less common elements of caregiver burden and family cancer history can fit within the standard CDM structure. This design will be documented so other digital health platforms facing similar challenges can learn from our approach.

The second deliverable is a comprehensive concept mapping specification; each item from our PRO instruments will be mapped to standard concepts in the OHDSI vocabularies. Where existing vocabularies fall short, particularly for caregiver burden concepts we will propose extensions to the OHDSI community for review.

Third, we will produce a quality validation framework with PRO-specific thresholds. Standard data quality metrics were designed for clinical data; we will adapt them for the unique characteristics of patient-reported data collected continuously between clinic visits.

Finally, we will contribute our implementation as a showcase case for OMOP 4 Sweden!, demonstrating that small digital health companies can participate meaningfully in Sweden's health data standardization efforts.

We plan to make quality-validated OMOP data available for OHDSI network queries, ready to contribute to federated research, by the end of 2026.

Conclusion

This project addresses three gaps that have limited the OHDSI ecosystem's ability to capture the full patient experience: the absence of prospective symptom data collected between clinic visits, the systematic invisibility of caregivers in observational research, and the underrepresentation of hereditary cancer risk in standardized databases. By demonstrating how a small digital health company can contribute research-grade PRO data to the community, we hope to encourage similar initiatives across Europe.

We invite feedback from OHDSI colleagues on our approach, particularly regarding vocabulary extensions for caregiver burden concepts and quality thresholds appropriate for continuously collected patient-reported data. We also seek collaborations with researchers interested in prospective cancer outcomes, caregiver wellbeing, and hereditary risk.

Oncoly began from a simple conviction, rooted in lived experience with cancer's impact on families: patients and caregivers should never have to pay for tools that support their care. By creating research value from the data they share, we can build platforms that serve patients first, and sustain themselves by advancing science, not by charging those who are already carrying enough.

References

1. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol. 2016;34(6):557-65. doi:10.1200/JCO.2015.63.0830
2. Di Maio M, Gallo C, Leighl NB, et al. Symptomatic toxicities experienced during anticancer treatment: agreement between patient and physician reporting in three randomized trials. J Clin Oncol. 2015;33(8):910-15. doi:10.1200/JCO.2014.57.9334
3. van Sandijk S, Prinsen P, van Hemelrijck M, et al. The manifold presentations of PROMs and questionnaires: patient-reported outcomes in OMOP use cases. OHDSI Global Symposium; 2022. [poster]
4. Wang L, Wen A, Fu S, et al. A scoping review of OMOP CDM adoption for cancer research using real world data. npj Digit Med. 2025;8:189. doi:10.1038/s41746-025-01581-7
5. Kühnel MB, Ramsenthaler C, Bausewein C, et al. Validation of two short versions of the Zarit Burden Interview in the palliative care setting. Support Care Cancer. 2020;28(11):5185-93. doi:10.1007/s00520-019-05288-w
6. Longacre ML, Applebaum AJ, Baird P, Miller MF, Golant M, Northouse L. Reducing informal caregiver burden in cancer: evidence-based programs in practice. Transl Behav Med. 2018;8(2):145–55.
7. Del Fiol G, Livnat Y, Goel P, et al. Identification of individuals with hereditary cancer risk through multiple data sources: a population-based method using the GARDE platform and the Utah Population Database. JCO Clin Cancer Inform. 2024;8:e2400142.
8. Swelife. This is Swelife. (*https://swelife.se/en/what-is-swelife/)* Lund: Swelife; 2025 [cited 2026 Feb 6]. Available from: https://swelife.se/en/what-is-swelife/