Connecting primary and secondary use of data in the EHDS: which benefits for the Long COVID use case?

The workshop titled *"Connecting primary and secondary use of data in the EHDS: creating the conditions for real data fluidity"* offers a rich, multi-perspective exploration of long COVID as both a clinical and data governance challenge. Through a sequence of interventions—from patients, clinicians, researchers, and policymakers—the discussion reveals not only the complexity of the condition itself but also the structural fragmentation that currently limits effective responses. What emerges is not a single narrative but a layered account of misalignment between lived experience, clinical practice, research infrastructures, and policy coordination.

The session opens with a framing scenario presented by **Stefano Dalmiani**, who introduces a fictional yet representative patient, Natalie, to illustrate the systemic gaps in managing long COVID. This narrative sets the tone by emphasizing the disconnect between symptom experience and diagnostic clarity. The description of a patient who "alternates between walking and taking sick leave" captures the fluctuating and poorly understood nature of the condition. Importantly, the scenario also introduces the idea of patient agency in data sharing, suggesting that patients are not merely subjects of care but potential contributors to knowledge production. This framing anticipates many of the tensions that later speakers will articulate more concretely.

Carmen Van Roy, speaking from direct patient experience, provides the most vivid and emotionally grounded account of the challenges. Her testimony is marked by both detail and systemic critique. She recounts a trajectory involving "75 doctors appointments in 12 hospitals and with 18 specialists," highlighting not only the burden of illness but the fragmentation of care. Her observation that "each specialist looked at one or two symptoms, even though this is clearly a multi system disease" encapsulates a core failure of current healthcare structures to address complex, systemic conditions. Her narrative also foregrounds the epistemic role of patients: "As a patient, I am connecting

the dots," she states, pointing to the paradox that those most affected are also those forced to integrate disparate pieces of information. The lack of a "single shared overview" of her data becomes both a practical and symbolic representation of system failure. At the same time, she introduces a forward-looking vision in which patients are "not just a patient… but also a source of knowledge," advocating for data systems that recognize and return value to contributors.

Robert Van de Stichele from I~HD and University of Gent offers a contrasting yet complementary perspective grounded in clinical practice and bottom-up research. His account of a single general practitioner building a cohort of over 400 patients demonstrates the potential of structured, consistent data collection even within limited institutional frameworks. He emphasizes the integration of diverse data types—coded clinical records, patient narratives, and biological samples—arguing that this constitutes "gold standard clinical data." However, his intervention also reveals tensions between data richness and interoperability. While efforts were made to convert records into standardized formats such as the International Patient Summary, he notes that this process leads to "loss of context," suggesting that standardization may come at the expense of clinical nuance. His insistence on multi-terminology approaches, including the need for human phenotype ontology beyond SNOMED, reflects a deeper concern that existing standards are insufficient for capturing the complexity of long COVID. His perspective implicitly challenges top-down harmonization efforts by demonstrating the value—and fragility—of locally generated, context-rich data.

Bianca Boxma-de Klerk introduces a more institutionalized and coordinated model through the Dutch Post-COVID Network. Her presentation reflects a high degree of systematization, with clearly defined work packages, metadata structures, and governance frameworks. The scale of the initiative—16,000 patients registered in a national portal—illustrates what can be achieved through coordinated national efforts. Yet even within this structured environment, nuances emerge. She acknowledges the tension between comprehensive data collection and patient burden, noting efforts "to limit the burden of filling out questionnaires." This reflects a recurring theme: the need to balance data ambition with human feasibility. Her emphasis on harmonization through common ontologies and codebooks aligns with broader European goals, but her admission that communication efforts still do not "reach everyone we would like to" underscores the persistent gap between system design and real-world engagement.

Helena Lira from the University of Helsinki's contribution shifts the focus toward high-dimensional, multi-modal research and the challenges of integrating advanced data types across borders. Her account of an EU-funded project involving omics data, AI analysis, and multiple countries highlights both the promise and the fragility of large-scale research infrastructures. The collapse of a private data platform provider ("neuromedia went to bankruptcy") serves as a cautionary example of dependency risks

in centralized systems. Her experience also illustrates ***the evolving nature of data governance***, particularly ***the need to revisit consent frameworks as research questions and technological possibilities expand***. Notably, she reflects critically on the limits of data itself, stating that "it's not about the data only, it's about best research questions." This introduces an important nuance: even the most sophisticated data infrastructures cannot compensate for unclear or misaligned research objectives. Her acknowledgment that biomarker discovery has so far yielded "weak biological signatures" further tempers expectations about purely data-driven solutions.

Stefan Schreck, representing the European Commission (DG-Santé), situates these diverse efforts within a broader policy context. His intervention emphasizes the structural constraints of EU action, particularly the primacy of member states in health system organization. He frames long COVID as a "cross-cutting public health and societal issue," extending beyond clinical care into education, employment, and social protection. His description of the Network of Expertise on Long COVID (NELC) as a platform for "structured collaboration" reflects an attempt to bridge fragmentation through coordination rather than centralization. However, his remarks also reveal institutional limitations. The reliance on member state needs to guide EU action introduces a reactive dynamic, and his acknowledgment that efforts to define a common case definition have been "not very successful so far" highlights the difficulty of achieving consensus even on foundational issues. His candid response to questions about project coordination—citing limited administrative capacity—adds a layer of realism to the discussion of European integration.

What are emerging cross-cutting issues ?

Across these contributions, several cross-cutting issues emerge. The most prominent is **fragmentation**—of data, care pathways, research efforts, and governance structures. Carmen's lived experience of dispersed data finds an echo in Robert's concerns about loss of context in standardized formats and in Bianca's efforts to harmonize disparate datasets. This fragmentation **is not merely technical but epistemic**: different actors hold partial, often incompatible views of the same phenomenon. As one participant notes, projects often "feel quite isolated," despite existing networks.

A second recurring theme is the **tension between standardization and complexity**. While harmonization is widely recognized as necessary for data sharing and comparability, multiple speakers caution against oversimplification. Robert's insistence that "SNOMED will not be enough" and Helena's observation of heterogeneous biological signals both point to the risk that standard frameworks may fail to capture the nuances of long COVID. This tension is further reflected in the difficulty of establishing a case definition, which remains elusive precisely because of the condition's variability.

Patient involvement emerges as both a principle and a challenge. Carmen's call for control over data and meaningful feedback—"if my data is used, I would like to receive

something back”—resonates with Bianca’s design of patient portals and Stefan’s emphasis on stakeholder engagement. Yet the extent to which patients are truly integrated into decision-making processes remains uneven. The idea of patients as “active contributors to knowledge” is widely endorsed but not yet fully operationalized.

Another important cross-cutting issue is the relationship between data and action. While the workshop is framed around data fluidity, several speakers question whether improved data alone will lead to better outcomes. Helena’s focus on interventions and Stefan’s emphasis on policy uptake both suggest that the translation from data to practice is far from automatic. The repeated concern that research results remain underutilized—“not self-evident that research result are actually taken up”—points to a persistent implementation gap.

Finally, the discussion highlights the temporal dimension of learning. Long COVID is presented both as a novel challenge and as part of a broader category of post-infectious conditions. Multiple participants note parallels with diseases such as ME/CFS and Lyme disease, suggesting that current efforts risk repeating past patterns of neglect. Carmen’s warning that “long covid is not new” and that similar conditions “didn’t receive enough funding in the past” serves as a reminder that the **current momentum must be sustained and broadened**.

In conclusion, the workshop reveals a field in transition. Significant advances have been made in data collection, infrastructure, and collaboration, yet these are unevenly distributed and often disconnected**. The challenge moving forward is not only to build more data systems but to align them with clinical realities, patient experiences, and policy needs**.

As one participant aptly summarizes, the goal is to “turn scattered experience into shared learning,” a task that requires not only technical solutions but sustained institutional and cultural change.