**Facilitating the spread of cancer (data)**

Kajsa Møllersen, Associate professor, Biostatistics, UiT - The Arctic University of Norway ([kajsa.mollersen@uit.no](mailto:kajsa.mollersen@uit.no))

**Background:** Cancer research relies on access to valuable and expensive data. Valuable in the sense that important findings can be retrieved, expensive in the sense that both time and effort must be in place to collect the data. This poster-talk presents the unfinished process of preparing a breast tumour data set for research and, most importantly, public sharing with other researchers. The specific data set consists of H&E slides, IHC, TMAs, and a number of other clinical and non-clinical data. The data is part of the larger Norwegian Women and Cancer (NOWAC) study, a nation-wide, multifaceted, epidemiological study - started in the 1990s and still going strong.

**Motivation:** The urge to spread cancer originated from a PhD project (Masoud Tafavvoghi) in machine learning on H&E slides. The local data set (n = 400) is not large enough, and Tafavvoghi had to access other data bases. Since we now are benefitting from other's willingness to share, why not contribute ourselves?

**Methods:** Quality control, digitization and anonymisation are some first crucial steps to data sharing. Ahead awaits structuring, labelling, metadata, storage facilities, and much, much more. This poster-talk describes some of the finished steps, and some of the steps that still lay ahead, some choices that have been made, and some mandatory requirements.

**Discussion:** We hope to invite to a conversation about data sharing, both on the practical level and on the more overall levels, from a research perspective, a health care perspective, and the patient perspective.

**Take home message:** Share your data!

Preference of presentation: Poster