Large-Scale Lung Function Analyses using an automated Data Stream: preliminary data from the SwissPedLung Data Stream

Xenia Bovermann[1,2], Anna Hartung[3], Léa Ho Dac[1], Tabitha Arn-Roth[1], Fabiën Belle[4], Claudia Kühni[1,4], Philipp Latzin[1]

[1]Division of Pediatric Respiratory Medicine and Allergology, Department of Pediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland; [2]Graduate School for Health Sciences, University of Bern, Switzerland; [3]Insel Data Science Center, Department of technology and innovation, Inselspital, Bern University Hospital, University of Bern, Switzerland; [4]Institute of Social and Preventive Medicine, University of Bern, Switzerland; contact: xenia.bovermann@insel.ch or philipp.latzin@insel.ch

Background 1 – SwissPedLung

SwissPedLung (SPL), a national data stream the national data stream "SwissPedHealth (SPH), envisions making **routine clinical lung function data available for research**. Its aims are:

- Combining data from lung function devices and electronic health records (EHRs).
- Collecting data from multiple pediatric pulmonology divisions in Switzerland.
- Include all patients under the age of 20 without denial of general consent between 2017 to 2023

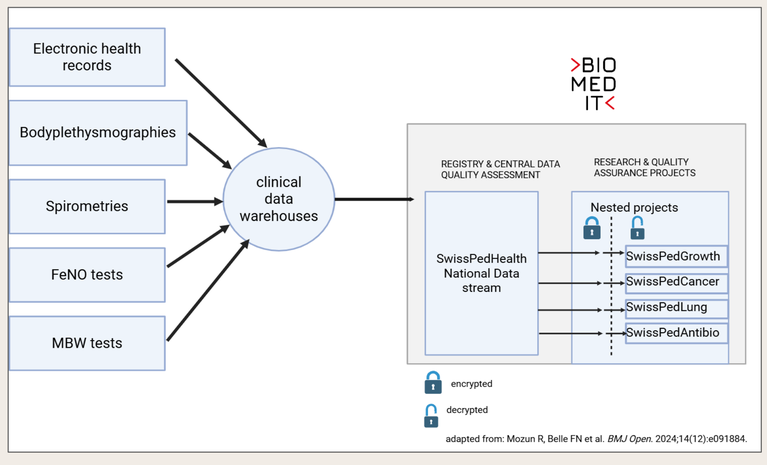

Figure 1: Simplified illustration of the data stream of SwissPedLung within SwissPedHealth. Lung function data from multiple machines with different software and data formats is combined in local clinical data warehouses (CDWs) with electronic health record data. After an encryption, the data is forwarded from the local CDWs to a central data platform, BiomedIT. Certified users can access the data platform to analyze their data. Figure adapted from [1]

Background 2 – Bronchodilator response

Guidelines on pediatric **asthma diagnosis** differ in their definition of a positive bronchodilator response (BDR):

- **ERS/ ATS** technical standard: increase in **FEV1 of ≥10%** relative to the **predicted** value (PV)) [1]
- **GINA** guidelines: ≥12% relative to the baseline value (BV) [2]

The impact of these differences is yet unkown.

Guideline	Positive BDR cutoff	Change
ERS/ ATS technical standard, 2022 [2]	≥ **10%** increase in FEV1 or FVC	Relative to **predicted** value
GINA, 2024 [3]	≥ **12%** increase in FEV1	Relative to **baseline** value

Table 1: Definition of positive BDR depending on guideline.

Aims:

- Building a national data stream for lung function data
- Using the data to analyse the impact of differences in BDR definitions

Results - SwissPedLung

The SPL data stream consists of interconnected concepts, structuring the data . Different data sources are linked to the concepts. Only data from one center could be analysed for now.

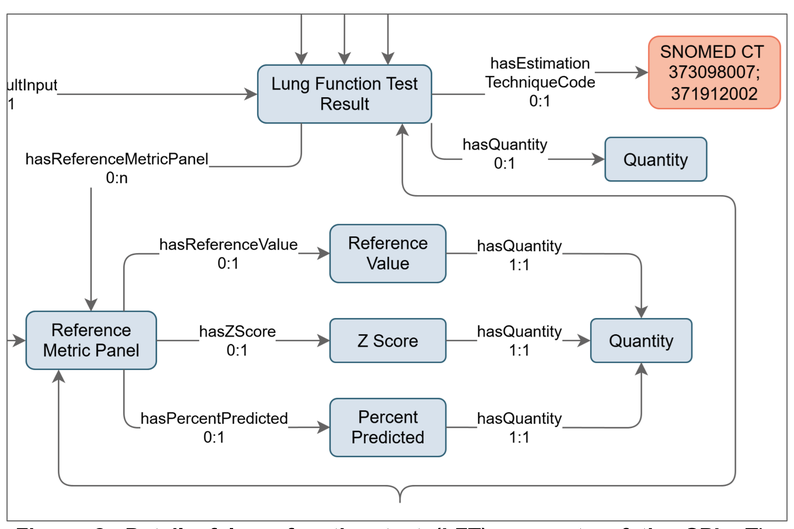

Figure 2: Detail of lung function test (LFT) concepts of the SPL. The concept "LFT results" includes reference values, z-scores and percent predicted. It is interconnected with numerous other concepts within the data stream.

characteristic	result
number of patients	4428
Number of spirometry tests (total) **Excluded (missing consent)** **Excluded (unacc. quality)**	14124 1792 780
females (%)	1874 (42,3 %)
age mean ± sd in years **(min; max)**	10.6 ± 3.6 (2.8; 20)
FEV1 BL mean ± sd (in %pred)	93.0±15.5
FEV1 BL mean ± sd (in zscore)	-0.6±1.3
mean FEV1 diff. pre to post ± sd	6.97±7.76

Table 2: Population characteristics of one centre of the SPL. N= 4285 for the FEV1. diff.
Z-Scores FVC and FEV1 according to [1]. Unacc. = unacceptable; FEV1 diff. = Difference in the volume exhaled in the first second between Pre and Post bronchodilation, BL = Baseline value

Results – bronchodilator response

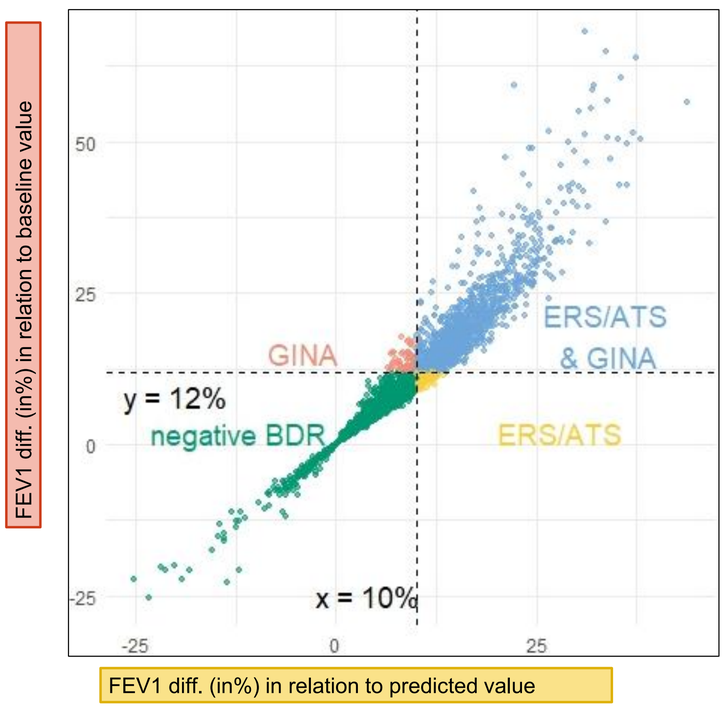

Figure 3: Plot of pre to post FEV1 differences. Calculated in relation to predicted (x-axis) versus baseline (y-axis) values. FEV1 diff. = Difference in the volume exhaled in the first second between Pre and Post bronchodilation. "GINA" and "ERS/ ATS" relate to the respective guidelines (see Background 2)

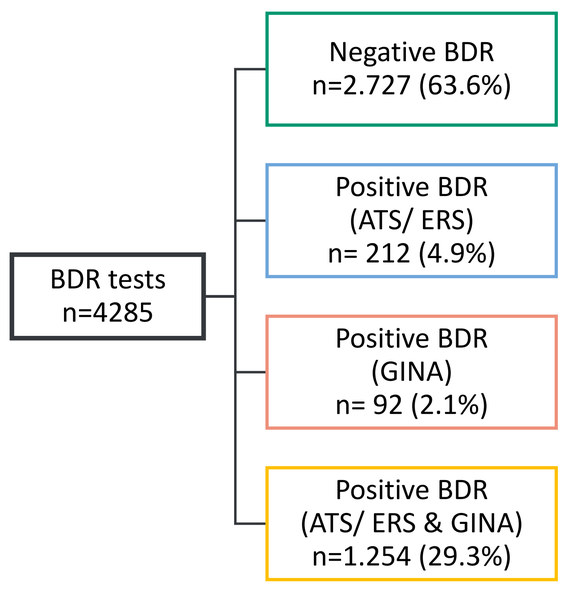

Figure 4: Number (%) of tests per BDR outcome group. BDR = bronchodilator response. "GINA" and "ERS/ ATS" relate to the respective guidelines (see Background 2)

Conclusions

1. Automated extraction of routinely collected lung function data and its combination with clinical data is feasible.
2. Highly complex data infrastructures may pose challenges for efficient scientific workflows.
3. Large-scale data enables detailed analysis of research questions.
4. Applying different guidelines for BDR calculation affects 7% of tests in our division

References
[1] Mozun, R. et al., BMJ Open, 2024 *BMJ Open, 2024. 14(12): p. e091884.*
[2] Stanojevic, S., et al., *Eur Respir J, 2022. 60(1).*
[3] Global initiative for Asthma. Main report. 2024
[4] Quanjer, P.H., et al., *Eur Respir J, 2012. 40(6): p. 1324-43.*