This document contains further information about our visualization prototype for lung function data. We provide additional screenshots of the interface, describe our study procedure in more detail and present some direct feedback from the participants.

Appendix A: Prototype

Figure 1 shows a screenshot of our prototype. While the spirometry measurements are visible in the image, others are not, but they could be accessed through scrolling down (e.g. body plethysmography measurements). In contrast to the previous image, Figure 2 shows a patient with more medical exams and therefore more visualized data points. This example nicely demonstrates the problem of having *Data-Dependent Horizontal Width* as mentioned in the results section (3.2). Depending on the screen width, the user is forced to scroll horizontally to analyze all data presented in the slope charts. Taking up so much space can also lead to problems when comparing values of different slope charts directly.

Appendix B: Requirement Analysis Workshop

This section briefly describes the procedure and results of our requirement analysis workshop. After preliminary discussions with one expert in the field of chronic lung diseases, we collected common tasks and potential solutions in a workshop with five pneumologists and two visualization experts. The aim of the workshop was to collect and classify tasks they routinely perform in their everyday work, and to gather typical questions which arise during treatment of a patient.

Due to the Coronavirus pandemic we carried out the workshop remotely, using a video conference and a collaborative whiteboard system. We prepared different workshop activities. After a brief introduction, we asked the participants to fill out post-it notes responding to one of three guiding questions:

- What do you want to know?
- What do you want to do?
- What do you want to see?

As a next step, participants were asked to draft visualization mockups for the first prototype.

Based on the results of the expert workshop, we developed the visualization prototype through an iterative design process. We consecutively refined the interface by having frequent discussion with one physician. Finally, we presented the prototype to the group of participants from the original workshop as well as to additional experts in the field of chronic lung diseases.

Appendix C: Study Details

In this section we discuss the study procedure and results in more detail. Additionally we provide some direct comments made by study participants.

Procedure

Each interview consists of several phases, starting with an introduction to the study. This is followed by a demonstration of the prototype features by the interviewer. Then, the participant gets some time to explore the interface on their own. After this familiarization comes the main part of the study. We prepared two tasks, each with different patient data. In the first task, the participant has to answer six questions with low to medium difficulty. In the second task, they have to answer six questions with low to high difficulty. Finally, there is a small discussion that allows for feedback, followed by the participant filling out the system usability score (SUS) questionnaire.

The patient data was chosen by an expert from a set of ten real-world records, such that they represent typical cases with different characteristics. These data records were anonymized before importing them into our prototype. The task questions and solutions were created in collaboration with a domain expert, with the goal of designing tasks that are relevant to the target users' everyday working activities. All questions received a difficulty label and a degree of importance. Examples for this coding are listed in Table 1. The question catalog was reviewed by another domain expert in terms of its relevance and validity and an external usability expert approved our study procedure.

After conducting the interviews, a transcript of the recordings was created manually. These transcripts were evaluated and searched for reoccurring remarks and any direct and indirect feedback. Statements in the transcribed texts were manually clustered into matching topics. In addition, all participant answers were compared to the expert's sample answers.

Results

In general, participant feedback was positive. Two participants explicitly mentioned the time-oriented design as a useful feature, whereas a third participant favourably noted the practicality for everyday life. In the following, we list a selection of positive comments, in original and translated form, made during the study.

"In total, I think it's really really good. (...) In particular, the long-term trend is good, but of course also the individual pre- and post-data." (P1)

Original: *"Also ich finde es insgesamt sehr sehr gut. (...) Vor allem der längerfristige Verlauf ist gut, aber auch letztendlich natürlich die einzelnen Pre- und Post-Daten"* (P1)

"For the tool itself, I find it relatively clear at first glance to see the value progression over time. This is (...) actually made quite clear by the curve (...) with the color coding" (P3)

Original: *"Vom Tool an sich finde ich das auf den ersten Blick relativ übersichtlich, um dann die Werte im zeitlich Verlauf zu sehen. Das ist (...) von der Kurve (...) mit der farblichen Codierung eigentlich ganz schön anschaulich gemacht."* (P3)

"Overall I think it's a really cool thing. Certainly very suitable for everyday use and very relevant to everyday life and once you've done it a few times, I think it's really easy to find your way around." (P2)

Original: *"Finde ich insgesamt eine sehr coole Sache. Sicher sehr alltagstauglich und sehr alltagsrelevant und wenn man es dann ein zwei mal gemacht hat, glaub ich, findet man sich auch total easy zurecht."* (P2)

Difficulty	Question	Goal
low	What is the FEV1 value of the patient for his last medical exam?	Find and read values using the visualization.
medium	How did the obstruction develop over time?	Comparing multiple values and degrees of severity over time.
high	Did the findings improve in between time A and time B?	Interpretation of all available data.

Table 1: *Translated example questions with varying degree of difficulty used during the interviews of the study.*

The suitability of our design was also supported by the large number of correct answers participants gave. For nine out of twelve questions all participants came to the same conclusions. Especially for the more complex questions, it was necessary for the participants to interpret and analyse the data. This resulted in minor answer differences, but all were still within the expectation of the expert.

Finally the comments and feedback of the interview were analysed, sorted and grouped into multiple categories. Three of the created categories with corresponding original and translated quotes can be found here:

Identifying the Date of a Measurement

"(...) sometimes I needed to take a look, where exactly which date can be found, (...)"
Original: *"(...) ich musste immer son bisschen gucken, wo genau welches Datum jetzt ist, (...)"*

Overview Chart Usage

"And secondly I was not sure, for what exactly the curves at the top - in this case the number 0.45 FEV1 - how it is related to the measurements"
Original: *"Und als zweites mir war nicht ganz klar, wofür diese Kurven hier oben - also was die Ziffer 0.45 FEV1 - in welcher Korrelation die hier zu den Messwerten steht."*

High Need for Vertical Space

"(...) now I need to shortly try to get both on one page (...)"
Original: *"(...) jetzt muss ich einmal ganz kurz ein bisschen hier beides auf eine Seite kriegen (...)"*

System Usability Scale

The results of the System Usability Scale were extracted using the calculations by John Brooke and are presented in Table 2. The SUS values show that participants generally view the interface to be usable. Three of the four participants have very high values over 80, whereas the remaining participant is less enthusiastic but still close to the others.

	P1	P2	P3	P4*
SUS value	87,5	77,5	92,5	95

Table 2: *Results for the SUS questionnaire. (*Only SUS questionnaire available for this participant. Interview was not valid due to technical difficulties.)*
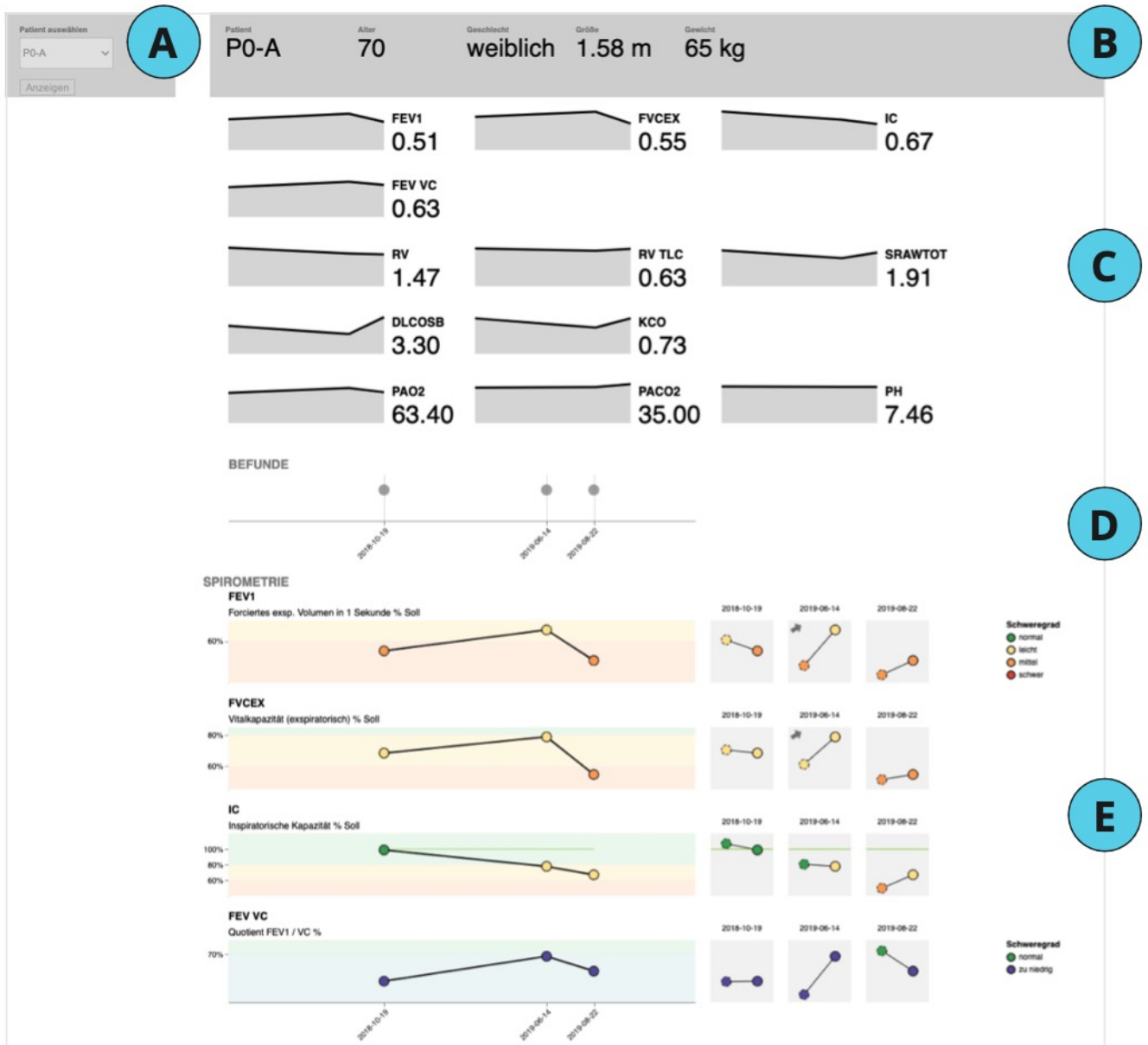
Figure 1: *An overview of the complete interface. Visible are the patient selector (A), the patient details (B), the overview charts (C), the timeline (D) and the line and slope charts (E). Regarding the line charts, four spirometry measurements are visible, while further measurements for example for body plethysmography are available but not shown.*
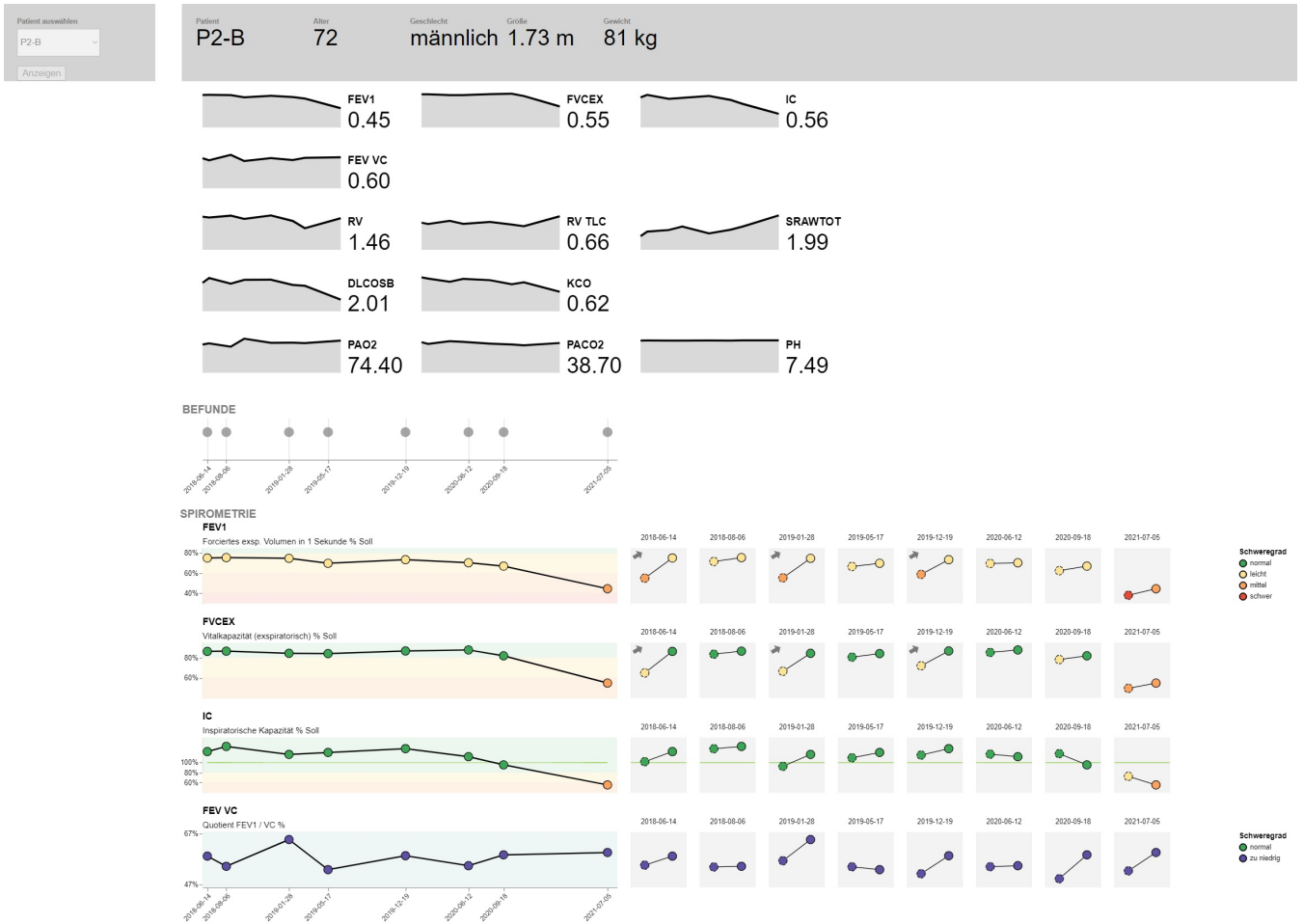
Figure 2: *Screenshot of the interface displaying a more extensive data set. The shown patient had eight exams, with the corresponding measurements presented in the line charts as well as the slope charts. In comparison to Figure 1 it is possible to detect the increased width necessary to display the complete interface.*