

BACKGROUND

The European Cystic Fibrosis Society Patient Registry (ECFSPR) collects demographic and clinical data from consenting people with CF in Europe. The Registry's database contains data of over 49,000 people living in 38 countries. The data from 21 countries is input manually into "ECFSTracker", the ECFSPR's custom-designed, data collection software. The remainder of the data comes from national registries, that collect data with their own data collection systems.

The data received by the ECFSPR is rigorously checked by inhouse statisticians, however, until 2018, no specific studies had ever examined the accuracy and consistency of the data at source. High quality data is essential for use in annual reports, epidemiological research and postauthorisation studies, therefore, a Data Quality Programme was introduced in 2018. It consists of four key elements, which are detailed below (item 4 is the main subject of the current publication):

1. A major update of ECFSTracker: improved, inbuilt data quality checks and controls, and an expanded set of data variables (launch software update: May 2019);
2. An evaluation and update of the organisation's Standard Operating Procedures Framework (2018 / 2019);
3. Liaison with the National Registries regarding data accuracy and consistency (2019);
4. Data validation visits to the centres in the countries that input data manually in ECFSTracker to validate the data at source (see table for results).

METHODS

A validation programme was introduced to quantify consistency and accuracy of data-input at source level, and verify that the informed consent – required to include data in the Registry – has been obtained in accordance with local and European legislation. Accuracy is defined as the proportion of values in the software that match the medical record, and consistency as definitions used by the centre that match those defined and required by the ECFSPR. The number of countries to validate: 20% of the total countries per year, max. 5 countries/year. In the selected country ≥10% of the centres are to be visited and 15-20% of patients' data validated. The visits are limited to centres with ≥50 patients.

In 2018 a team consisting of personnel from the ECFSPR and the Interdisciplinary Centre for Clinical Trials at the University Medical Center in Mainz, Germany, carried out a series of pilot visits in 4 countries. The team verified annual data for 2016, at source, in CF centres in Austria, Portugal, Slovakia and Switzerland.

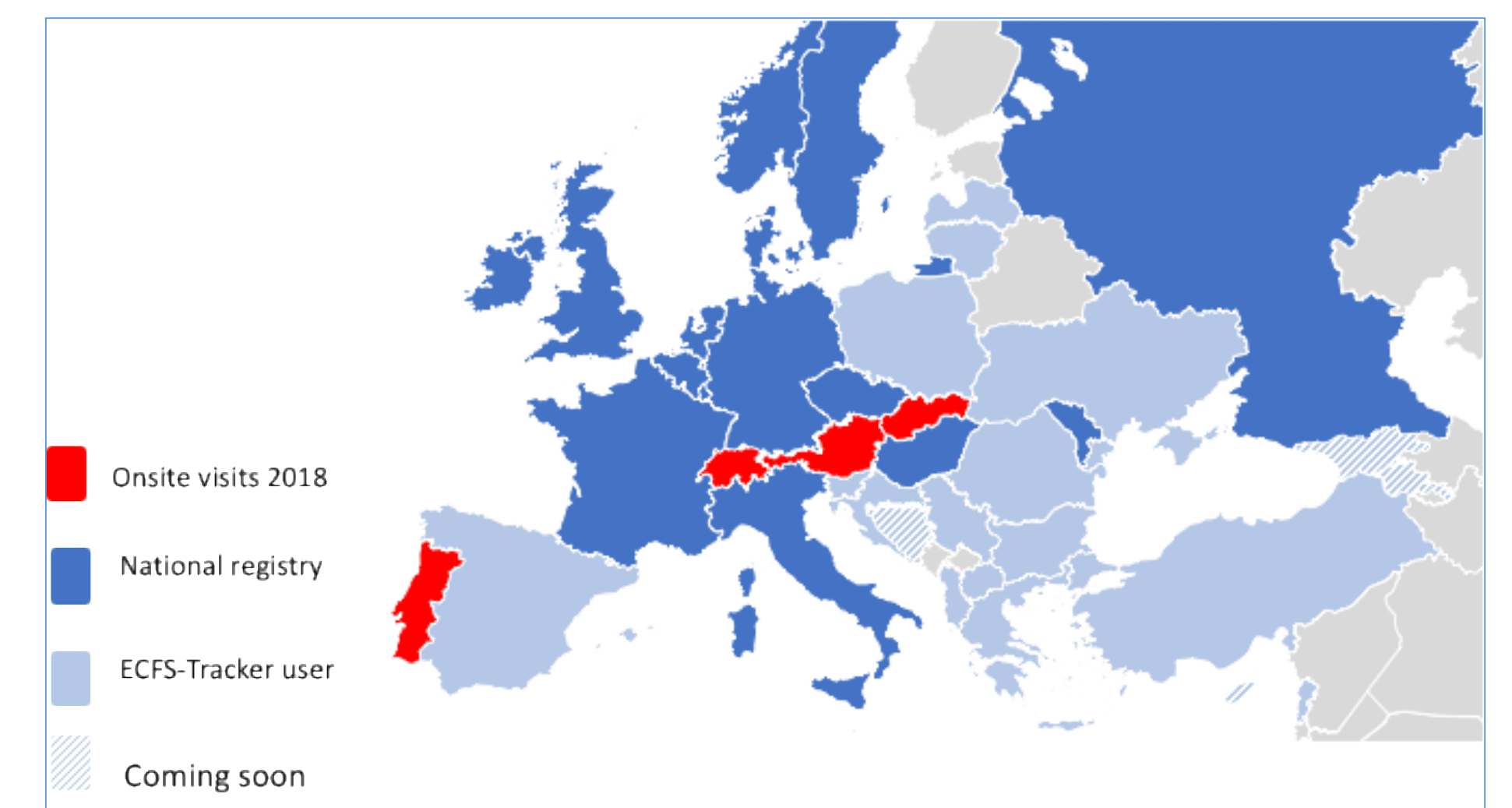
Selection of patients in each centre:

- Random selection of patients;
- Goal: 50% of adult patients, 40% of patients in the age group 6 - 17 years, 10% in the age group 0 - 5 years.

Selection of variables:

- Variables that are of particular relevance for key reports;
- Variables where inconsistencies or inaccuracies are suspected, or where input of correct data in the software has been highlighted as challenging by users;
- Demographic, diagnostic and transplant data, anthropometric and best lung function measurements, results of selected bacterial infections, medications and complications.

AIM: ≥95% of the data is correct.



RESULTS

In a one day visit the aim of the programme was explained to the centre, the data included in the Registry were compared with the medical records, the outcomes and recommendations discussed, and a final report provided to the centre. Challenges proved to be: the informed consent (re-consent at adult age or when the patient moved centre), mutation information (genetic laboratory report missing), and different interpretations of the definitions.

10 centres (24%), reporting ≥ 50% of all patients in their countries, were selected from the 41 centres in 4 countries (Austria, Portugal, Slovakia, Switzerland).

Demographic, diagnostic and transplant data was checked for 489 patients (21.0%*), and a selection of clinical data was checked for 463 patients (19.9%*).

*= the % of total patients in these countries.

Variable type	Variables verified	Correct (= accurate and consistent) Total (range for each country)	Comment
Demographics	Birth date (month and year only) Gender	98.8% (96.2 - 100%) 99.8% (99.5 - 100%)	
Genetic information	Mutation	77.4% (55.2 - 91.7%)	No source data 21.4% (4.1-44.5%); Incorrect data 0.9% (0.0-5.1%)
Transplantation	Organ (Lung, Liver) Year of transplant	99.8% (99.1 - 100%)	
Anthropometrics	Weight Height	92.2% (77.5 - 97.5%) 92.8% (81.6 - 97.5%)	Definition criteria in centre/s selected inconsistent with ECFSPR definition: "height and weight at best FEV1% pred of the year"
Lung function	Best FEV1% pred of the year	86.4% (38.8 - 92.6%)	Definition criteria in centre/s inconsistent with the ECFSPR definition
Medication	Inhaled antibiotics DNase Pancreatic enzyme use	96.1% (93.9 - 98.6%) 98.1% (96.3 - 99.3%) 97.6% (93.8 - 99.3%)	
Microbiology	Chronic Pseudomonas aeruginosa infection Chronic Burkholderia Spp infection	95.0% (85.7 - 99.3%) 97.0% (85.7 - 99.3%)	
Complications	Liver disease Major Haemoptysis Diabetes treated with daily insulin	86.8% (84.7 - 91.8%) 94.6% (86.4 - 100%) 97.2% (93.8 - 100%)	Uncertainty regarding the definition of "liver disease w/o cirrhosis"

Legend: Green > 95%, Blue 90-95% Red: < 90%

CONCLUSION AND NEXT STEPS

The ECFSPR dataset showed a high degree of accuracy and consistency for most of the data that was verified at source. To further improve performance it is recommended that CF centres use a reliable source for genetic information, adhere to the ECFSPR definition for best lung function (including anthropometry), and that the ECFSPR clarifies the definition of liver disease.

The validation visits delivered are essential to optimise data quality at source, ensure that centres are aware of the importance of correct informed consent, and encourage a dialogue to gain insight in how procedures, software, support and training can be improved. The visits gained insights into data quality and other related issues, both for the Registry as the centres, and highlighted that there is room for further improvement and consultation. The lessons learned are presented and discussed in Registry meetings and training.

ACKNOWLEDGEMENTS

The onsite validation visits would not have been possible without the commitment and support of the centre directors and their colleagues. We thank them for the excellent support and discussions:

Austria: Sabine Renner and Brigitte Mersi (Universitätsklinik für Kinder und Jugendheilkunde, Wien), Helmut Ellemunter and Johannes Eder (Medizinische Universität Innsbruck), Andreas Pflieger (Medizinische Universität Graz). Portugal: Luisa Pereira and Pilar Azevedo (Hospital Santa Maria, Centro Hospitalar Lisboa Norte), Fernanda Gamboa and Teresa Silva (Hospitais da Universidade de Coimbra). Slovakia: Hana Kayserova and Nina Bliznakova (Klinika detskej pneumologie, SZU UN Bratislava). Switzerland: Andreas Jung (Kinderspital Zürich), Philip Latzin and Romy Rodriguez (Inselspital Bern), Christian Benden and Thomas Kurowski (Universitätsspital Zürich), Reta Fischer and Patrizia Bevilacqua (Quartier Bleu, Bern).

A special thanks to the native language speaking volunteers who joined the discussions to help overcome language barriers: Milan Macek for Slovakia and Ricardo Vieira for Portugal, and to Nadine Wollscheid (Interdisciplinary Center for Clinical Trials, University Medical Center Mainz, Germany), who conducted the visits.

