



EUROPEAN MEDICINES AGENCY  
SCIENCE MEDICINES HEALTH

Thursday 29 November 2012  
EMA/769571/2012  
Human Medicines Development and Evaluation

## Report of the workshop on endpoints for cystic fibrosis clinical trials

European Medicines Agency, London, 27-28 September 2012

### Introduction

The Agency held a two-day workshop on 27-28 September 2012 on outcome measures for cystic fibrosis (CF) clinical trials.

Experts in the field of CF (N=12) together with European drug regulators (N=14), representatives from 24 pharmaceutical companies, and 1 patient representative (with the second one unfortunately prevented from participation because of an acute pulmonary exacerbation) discussed and compiled current scientific evidence on outcome measures for evaluating medicines targeting CF lung disease, bronchopulmonary infection and exocrine pancreatic insufficiency.

The workshop was also observed by FDA, PMDA and Health Canada via broadcasting.

The morning session of the first day was dedicated to set the scene for the break-out group discussions. The first speaker was Prof Stuart Elborn, current president of the [European Cystic Fibrosis Society \(ECFS\)](#) and director of an adult CF centre, giving an update on the changing demographics of patient population and evolving standards of care and the implications for clinical-trial design and endpoints.

Prof De Boeck, current vice-president of the European Cystic Fibrosis Society (ECFS) and first director of the [ECFS clinical trial network \(CTN\)](#) and director of a paediatric CF centre talked about "Difficulties/challenges encountered – look into the future: academia perspective."

The industry perspective on "Difficulties/challenges encountered – look into the future" was presented by David Waltz from Vertex Pharmaceuticals Incorporated, followed by the patient perspective presented by Emma Lake, [Cystic Fibrosis Trust, UK](#).

The regulatory perspectives addressed difficulties encountered by Scientific Advice Working Party ([SAWP](#)), Committee for Orphan Medicinal Products ([COMP](#)), the Paediatric Committee ([PDCO](#)) and the Committee for Medicinal Products for Human Use ([CHMP](#)) when evaluating drug applications of medicines for CF.

At the end of the morning session, Dr Siviero from the Italian Medicines Agency talked about "Endpoints for added clinical benefit in view of health technology assessment (HTA)."



Presentations are available on the [EMA webpage](#) together with this report.

To allow for sufficient in-depth discussion, the participants were then assigned to three break-out groups:

- Group 1: Pulmonary disease - co-chaired by Steffen Thirstrup, head of CHMP respiratory drafting group and Christiane De Boeck ;
- Group 2: Bronchopulmonary infection - co-chaired by Mair Powell, head CHMP anti-infective working party and Stuart Elborn;
- Group 3: Exocrine pancreatic insufficiency - co-chaired by Elmer Schabel, head CHMP gastroenterology drafting group and Ian Taminiau, PDCO member.

The three groups were tasked to discuss and answer the list of questions<sup>\*</sup>, prepared and circulated to all participants in advance of the meeting.

Answers to the questions should lay the ground for revising the current EMA [Guideline on the clinical development of medicinal products for the treatment of cystic fibrosis](#), and for establishing a “to do list” for academia and /or industry to obtain the lacking evidence needed to establish the validity of candidate surrogate endpoints.

## Group 1: Pulmonary disease

Topics discussed in Group 1 targeted:

1. Clinical endpoints.
2. Biomarker / Surrogate endpoints.
3. Patient related outcomes (PROs).

A list of core pulmonary and extra-pulmonary outcome measures was defined.

At present, FEV1, despite its major limitations, still remains an important outcome measure for clinical efficacy. However, promising new candidate outcome measures such as lung clearance index and computer tomography, particularly useful for detecting early lung disease, should be evaluated in parallel in order to establish their validity as surrogate endpoints.

Regarding patient related outcomes (PROs) it was generally agreed that the CFQ-R is currently the best validated instrument and should be used in clinical trials. However, it was proposed to start thinking of drastically simplifying PROs with a minimum number of questions only addressing the ultimate goal of any therapeutic intervention: whether or not the patient feels better after the intervention, whatever “better” might mean for the individual patient. This proposal would meet patients’ needs as the burden and reluctance to fill in lengthy questionnaires was expressed by the patient representative. Patients should be actively involved in the development of such a simplified instrument.

The detailed answers are attached in the [Annex I](#).

The group also [identified areas in need of further research and education](#).

---

<sup>\*</sup> See annexes to this report below.

## Group 2: Bronchopulmonary infection

Group 2 discussed the following four topics:

1. General microbiological issues, including questions related to sampling approach, sputum processing, methods for culture and assessment of reproducibility and sensitivity.
2. PK/PD studies – particular focus on inhaled antibiotics.
3. Antibiotic eradication treatment.
4. Chronic suppressive antibiotic treatment with particular focus on Pseudomonas infections.
5. The detailed answers are attached in the [Annex II](#) together with list of [identified areas in need of further research](#).

## Group 3: Exocrine pancreatic insufficiency

The discussion in Group 3 focused on two areas:

1. Pancreatic enzyme replacement therapy.
2. CFTR modifiers and their impact on GI outcomes/endpoints.

The detailed answers are attached in the [Annex III](#) together with list of [identified areas in need of further research](#).

## Outcomes

The outcome of the breakout group discussions were reported back to the plenary on the afternoon of the second workshop day. The conclusions of the general discussion were that there is a need for:

- new outcome measures;
- novel trial designs to accommodate for the limited patient population and the increasing drug development pipeline in CF;
- definition of endpoints for added clinical benefit in view of HTA and to maintain orphan status.

In addition, existing international (North American, European) and national CF registries should be used more widely as they can provide data:

- to compare the disease course in differing patient groups to identify optimal patient cohorts for interventional studies;
- to describe variability of outcome parameters without interventions in different age groups and perform power calculations;
- for feasibility assessment;
- for pharmacovigilance studies.

Summarising the discussions and the conclusions, it was agreed that there is need to revise the EMA CF guideline.

## Agreed action points

The report together with its annex to be sent to the CHMP Respiratory Drafting Group as well as the CHMP Gastroenterology Drafting Group for discussion in order to initiate revision of the CF guideline.

# Annex I to the workshop on endpoints for cystic fibrosis clinical trials

## Breakout group 1: Pulmonary disease

### **Background:**

- Lung disease major determinant for morbidity and mortality in CF.
- Advances in symptomatic treatment options and patient care resulted in great improvement of survival and changes of the demographics of the CF population.
- Majority of children with CF have spirometry within the normal range and rarely experience pulmonary exacerbations.
- Statistically significant changes in pulmonary function and exacerbation rate become more difficult to demonstrate as baseline population values improve.
- Lower event rates translate into larger sample sizes needed to detect smaller, yet still meaningful effect sizes.
- For the first time disease modifying drugs are available, which should ideally be administered prior to development of lung damage.
- FEV1 as surrogate for mortality usually required by EMA as primary endpoint for most pulmonary indications.
- Spirometry (FEV1) unable to detect early lung damage, early bronchiectasis and cannot be measured in pre-school children.
- An ever increasing drug development pipeline adds to challenge of conducting clinical trials in a limited patient population: According to the clinicaltrials.gov website, there were a total of 143 clinical trials actively enrolling CF patients as of April 2012.
- The need for large clinical trials of one product may hinder the enrolment of patients into other trials, thus affecting opportunities to study another product for the same disease condition.

### **Questions - Clinical endpoints:**

- 1. What is the more relevant clinical outcome from a patient point of view: Frequency (total number) of pulmonary exacerbations OR time to event OR need for hospitalisation? Other?**

Time to exacerbation is the most feasible endpoint.

Treatment for pulmonary exacerbations and hospitalisations and are both important for patients with CF. Since the need for hospitalization is subjective (clinician and patient dependent with different rates of home IV use between centres) they are likely not independently useful. The time to exacerbation and frequency of exacerbation are linked; shorter time intervals between exacerbations are a risk factor for lung function decline (Waters V et al., Eur Respir J 2011). Longer study duration is needed to calculate frequency of exacerbations in a reliable way. Therefore, time to pulmonary exacerbation is the most feasible endpoint. The frequency of pulmonary exacerbations increases with age and with

lung disease severity (Goss thorax). In all age categories, it is useful to record the history and especially the of numbers of IV treated pulmonary exacerbations in the year prior to entry into the trial.

## **2. What could be used as universally agreed definition for acute exacerbation in regulatory trials?**

The definition of Fuchs in its original form (4/16 symptoms leading to IV antibiotic treatment) or in its modified form (4/16 symptoms leading to any change in antibiotic therapy) can be used as universally agreed definition of acute exacerbation. For consistency, all symptoms included in the definition of a pulmonary exacerbation should be monitored at every trial visit.

The frequency of pulmonary exacerbations is linked to lung disease severity (Goss , Thorax). Pulmonary exacerbations requiring intravenous antibiotics/hospitalisation are linked to lung function decline, bronchiectasis, and mortality (Meyer-Hamblett 2002, AJRCCM; Konstan 2007, J Pediatr; Sanders 2010, AJRCCM; Waters 2011, Eur Respir J). This relationship is less clear for milder exacerbations that are usually treated with oral antibiotics (Definitions of pulmonary exacerbations have not been validated a priori and none of the definitions capture severity of the event. The most sensitive symptoms to define a pulmonary exacerbation depend on the population studied (Rabin Pediatr Pulmonol 2004). So far, the Fuchs criteria used in the phase III trial of dornase alfa (symptom plus intervention defined) have been used most often. (Fuchs 1994, NEJM).

In a population that has established CF lung disease one should also separately analyse increase in symptoms requiring intravenous antibiotics in hospital or at home since these are clearly linked to lung function decline. In patients with CF and mild CF lung disease, intravenously treated exacerbations are rare and therefore not a feasible outcome measure.

It is likely that the definition used is not that important, but that –in analogy with the COPD and bronchiectasis trials- the use of new/more antibiotics is the key part. The simplified definition proposed by the European Cystic Fibrosis Society consensus (Bilton 2011, J Cyst Fibros) can be used in parallel to evaluate its sensitivity: need for additional antibiotic treatment as indicated by a recent change in at least 2 of the following:

For health technology assessment (HTA) not only the event but also the type (e.g. hospitalisation versus treatment at home) and duration of the intervention should be measured.

## **3. Is it possible to define standardized criteria when to initiate systemic antibiotic treatment?**

It is possible, but aiming for it will decrease study feasibility.

Practice patterns differ with physicians both treating events that will not fulfill the criteria and decide not to treat even though the definition is met. In a multicentre phase 3 study, randomisation will most likely balance such differences out between centres.

Aiming for standardized criteria to initiate systemic antibiotic treatment will greatly impact on study feasibility: physicians and patients may consider it unethical to change their routine practice for the purpose of a clinical trial.

## **4. Defining a core outcome set of pulmonary and non-pulmonary endpoints, to be used in phase III trials, seems warranted to make results of several trials comparable. The primary endpoint may then be chosen among them, or be a composite endpoint, or two co-primaries, as long as the core outcome set is included.**

Agreed.

This would facilitate comparing data from different trials. We however need to point out that in most phase III trials to date, treatment response was heterogeneous to different outcomes; i.e. lung function response does not predict benefit for the pulmonary exacerbation outcome parameter (e.g. Saiman et al, AJRCCM 2005). Thus, a combination of endpoints may not be better to define responders versus non-responders. In case of a composite endpoint or co-primary endpoints the impact on statistical power needs to be considered. New modelling techniques will explore the potential of prediction models to define responders.

The list of core outcomes from which the primary outcomes for a specific study should be chosen includes:

- Lung function: FEV1 and in early disease FEF 25-75, LCI;
- Pulmonary exacerbations / need for additional antibiotic therapy/ hospitalisation;
- PRO measures;
- Imaging: chest CT score for bronchiectasis/air trapping;
- Biometry: Weight and in children also Height;
- Biomarkers depending on the drug class and the supposed mechanism of action;
- Safety outcome measures appropriate for the compound under study.

#### **5. How could the European CF registry be used to support evidence generation in CF?**

The European CF patient Registry (N=25 000) contains core demographic data, genotype, FEV1, microbiology parameters and qualitative data on specific complications. The data set of the centres working in the Clinical Trial Network (N=13000) contains also more specific patient characteristics and is in many centres encounter based. These rich data resources could be useful for the following assessments:

- 5.1. Describe the current disease course and define important medical needs.
- 5.2. Compare the disease course in differing patient groups to identify optimal patient cohorts for interventional studies
- 5.3. Describe variability of outcome parameters without interventions in different age groups and perform power calculations
- 5.4. Feasibility; impact of proposed inclusion/ exclusion criteria on possible clinical trial recruitment
- 5.5. Data modelling techniques to explore alternative clinical trial design
- 5.6. Phase 4 pharmacovigilance studies for safety assessment of licensed drugs
- 5.7. Pharmaco-economic data such as (potential) drug use and cost

#### **6. What additional endpoints acceptable for HTA should be included as secondary/additional endpoints in regulatory trials to assess real life benefit?**

- 6.1. PROMs, Validated Quality of life tools including treatment burden;
- 6.2. Time off school/ work/ carers work;
- 6.3. Use and duration of additional courses of drugs (incl antibiotics);
- 6.4. Impact on resource use, health economics/ cost savings;

#### 6.5. Prevention of CF specific complications.

It is suggested to consider need to add HTA metrics at the time of study design, i.e. making phase III trials useful for both regulatory and HTA decisions

### **Questions - Biomarker / Surrogate endpoints:**

For most biomarkers and surrogate outcomes we do not have sufficient information on the natural variability in patients with CF. For many biomarkers this variability most likely differs according to patient age and lung disease severity. Knowing the background variability of a parameter will help to interpret the significance of the mean effect size of an intervention. MCID is not easy to determine and not easy to measure. Knowing an MCID for a measurement is probably most useful for responder analysis.

#### **7. What effect size of FEV1 improvement can be considered to be clinically relevant? Is there a minimal important difference defined?**

As FEV1 is linked to mortality, any significant difference between placebo and active treatment is potentially clinically relevant. Not only the improvement in FEV1 but also the sustainability of the improvement should be taken into account. Changing the FEV1 rate of decline would be the most meaningful effect but this is not feasible in practice since it would require a long study duration and/or a large sample size (Konstan, JCF 2010). A treatment effect equivalent to the average annual loss in FEV1 can be considered as clinically relevant.

A MCID for FEV1 has not been defined. This value would help to define treatment responders. We indeed have insufficient information about the short term variability of FEV1 in (subsets of) patients with CF. In an old study (Cooper 1990, *Pediatr Pulmonol*) the median FEV1 CV was 5.8%. In a recent Danish registry study (Taylor-Robinson 2012, *Thorax*) the short-term variability of FEV1 was 6.3%. The stability and treatment regimen of these patients was however not defined.

#### **8. When conducting non-inferiority trials, what non-inferiority margin (difference on FEV1) is acceptable to demonstrate non-inferiority (versus the clinically relevant difference)?**

Since we do not have a MCID for FEV1 we do not have robust data to set this limit. This implies there are inherent problems with non-inferiority trials in CF.

#### **9. What is the current yearly decline in FEV1 in patients not yet chronically infected with Pseudomonas versus those chronically infected?**

The yearly decline in FEV1 depends on the baseline characteristics of the population such as birth cohort, age, pancreatic status and baseline lung function. The rates of decline in lung function in young adults with CF diminish with successive birth cohorts and patients infected with *Pseudomonas* traditionally had a greater average decline in FEV1 (-1.6% v -1.1%) (Que 2006, *Thorax*). Konstan et al (JCF 2010) report mean rates of FEV1 change in a population of CF > 6 years and with baseline FEV1 above 70%. Median age specific year to year changes in FEV1 % predicted (Liou 2010, JCF) vary from close to 0% up to 4% during adolescence and young adulthood. Although the median changes are low, the variability is high, with 5 to 20% of the population having changes more than 10% predicted. A mean (SD) year to year change in FEV1 of only -1.22 (9.17) was also reported in the Belgian population (De Boeck ERJ 2011).

Chronic mucoid *Pseudomonas* infection has repeatedly been shown to be associated with faster rate of FEV1 decline (Ballmann 1998, *Thorax*; Parad 1999, *Infect Immun*; Li 2005, *JAMA*). Using Gaussian modelling, *Pseudomonas* infection was associated with a -0.51 (95%CL -0.72; -0.29) 'supplementary' lung function decline in the Danish CF population (Talyor-Robinson *Thorax* 2012). Some data analyses

support the fact that -in more recent years, the impact of chronic mucoid Pseudomonas infection may be smaller (Waters 2012, Eur Resp J; Dasenbrook 2008, AJRCCM). The annual decline in non-Pseudomonas infected patients is nowadays in many CF populations so low that mostly large patient numbers are required when FEV1 change is used as primary outcome measure.

**10. Are other surrogate lung function parameters available which would fulfil the requirements as listed on page 1 and would be capable to detect subtle improvement if baseline lung function is normal?**

In the original retrospective Canadian cohort linking lung function with mortality, not only decline in FEV1 but also rate of decline in FVC and FEF25-75 related to survival age. Compared to FEV1, the decline in FVC was less steep, that in FEF25-75 was steeper (Corey M, J Pediatr 1997). Both can therefore be used as additional parameters. They probably offer no real advantage since compared to FEV1, FVC has similar variability and FEF25-75 has increased variability with repeat measurements. FEF25-75 may be especially appropriate for use in patients with early lung disease (e.g. Quan 2001, J Pediatr; rhDNase early intervention trial). The same may be true for FEF75 which has been used successfully in a number of smaller studies aiming to reduce small airways disease (van der Giessen 2006, Pediatr Pulmonol; Bakker 2011, Eur Resp J). Though there is more variability in these peripheral flows the signal is larger. It is advised to evaluate whether the sensitivity of these measurements can be increased by using the mean of repeat measurements at baseline and at end of intervention.

**11. How could lung function be evaluated in pre-school children? Are lung function tests available which are feasible, standardised, reliably reproducible and feasible across several study sites for multicentre studies?**

Standardized and reliable lung function tests are available for use in the preschool age group. The feasibility depends on the expertise of the centre and is also dependent on the need for and side effects from sedation. The test sensitivity in this age group depends on the parameter used. In a multicentre observational study three different lung function techniques were compared in 93 CF subjects and 87 controls 3-5 years old. Acceptability rates were lowest for spirometry (55%) and highest for inductive plethysmography (77%). Spirometry success increased with age and having previous acceptable measurement. FEV1, FEV0.5, and FEF25-75 were lower for CF subjects than for controls; spirometric z-scores declined with age. Inductive plethysmography measures of thoracoabdominal asynchrony were greater for CF subjects than for controls. Forced oscillation indices did not distinguish CF from controls. FEV1 and FEV0.5 are able to detect the smallest treatment effect for a given sample size (Kerby 2012, Pediatr Pulmonol).

**12. How to evaluate the treatment effect of CFTR modifiers in very young children with no or only little irreversible lung destruction?**

Proof of efficacy in this age category will depend on demonstration of prevention or decrease in progression of lung function abnormalities and structural damage. It will likely require study duration of at least 2 years.

Depending on the context (efficacy of the compound proven in other age groups; numbers of patients available) the effect of the compound on biomarkers such as sweat chloride that directly measure CFTR function, can be considered as intermediate proof of efficacy. However, the effect of the CFTR modifiers on the sweat gland may differ between drug classes (potentiator, corrector or read through drug) and depend on the tissue distribution of the compound and not necessarily reflects the effect on other organs.

Severity and extent of bronchiectasis and of trapped air as measured by bronchiectasis and trapped air CF-CT score can be used to quantify structural lung damage and its prevention by a CFTR modifier.

More accurate and precise image analysis methods are under development awaiting proof as a biomarker. If started very early on, CFTR modifiers have in theory the possibility to prevent complications like bronchiectasis.

CFQ-R is probably not very helpful due to a ceiling effect in this age group with mild lung disease.

Rate of pulmonary exacerbation (e.g. ISIS definition Rosenfeld M, JAMA 2012) can be used as an outcome parameter as are height and weight evolution, but both require a long follow up period.

Infant pulmonary function tests are not ready for use as primary outcome parameter in multicentre trials (Davis 2010, AJRCCM). The safety, feasibility and ability to detect abnormalities in infants with CF by infant lung function were assessed in this multicentre trial. Acceptable measurements could not be obtained in a substantial proportion of participants, especially at less experienced sites. Twelve % of participants withdrew due to side effects of sedation. Key lung function measurements were significantly different between CF and historical controls, but rather high patient numbers per treatment arm (N >75) are needed to detect a treatment effect of 7 to 17% of the baseline value, depending on the parameter chosen.

LCI has promise as outcome parameter in infants (Belessis 2012, AJRCCM) because it is higher in infants with CF compared to controls and is higher in subjects with Pseudomonas infection and increased inflammation. It has been used in an interventional trial in one interventional trial in this age group (ISIS trial).

For further details we refer to the summary of the Rotterdam workshop meeting minutes (Stick S et al., Eur Respir J 2012).

### **13. What is the current evidence regarding lung clearance index (LCI)? Has LCI been reproducibly shown:**

- **to detect even subtle treatment effects?**

LCI is more sensitive than FEV1 to detect disease in patients with relatively preserved lung function (Aurora 2004, Thorax) and has been linked to bronchiectasis (Gustafsson 2008, Thorax). LCI in preschool years predicts subsequent lung function (Aurora 2011, ACRCCM). LCI has detected treatment effects in 3 interventional studies in school age children; one of which was performed as a multi-centre trial (Amin 2010, Thorax; Amin 2011, Eur Respir J; Davies 2012, American Thoracic Society International Conference. San Francisco). A study performed as part of the ISIS trial (JAMA 2012) also suggests that LCI can detect a treatment effect in infants and preschool children (Subbarao et al., manuscript submitted).

- **that improvements translate into long-term patient benefit?**

Whether improvements in LCI translate into long term benefits is not yet known. To demonstrate this is particularly challenging in a population with early disease with few symptoms and no medium term mortality. LCI as well as changes in LCI during treatment in the 3 interventional studies mentioned above are correlated to changes in FEV1 making it likely that improvements in LCI will translate into long term benefit for patients. Also the fact that LCI in preschool years predicts later FEV1 supports this (Aurora 2011, ACRCCM).

- **that it can be reproducibly used in multicentre trials across all age ranges?**

The ivacaftor trial in patients with G551D and mild lung disease proves that LCI can be used in a multi-centre trial. Further evidence is needed especially in infants and young children (Davies 2012, American Thoracic Society International Conference. San Francisco). So far no multi-centre trial has been conducted in this age group and only one single centre study suggests its utility to detect a

treatment effect in preschool children and infants (ISIS trial; Subbarao et al., submitted). However, feasibility and reproducibility in this study as well as in studies in older children was high.

In conclusion, LCI will be especially useful for trials with patients with early lung disease since, compared to FEV<sub>1</sub>, smaller sample sizes are needed to detect treatment effects. It is ready for use in children from age 6 years on. It has great promise in the preschool and infant age group, but more insight in natural variability is needed before it can be used in multicentre trials in this age group. LCI (in its current set up) has less potential for trials in patients with advanced lung disease due to long measurement times and greater variability.

#### **14. Is a standardised methodology available?**

Standardization guidelines have been published (Robinson 2012, Eur Respir J) and a SOP is in development by the ECFS CTN (expected completion December 2012).

#### **15. What effect size of LCI improvement can be considered to be clinically relevant?**

The MCID for LCI has not been defined. Effect sizes in studies have ranged from -1 to -2 depending on the type of intervention and the duration of treatment (Amin 2010, Thorax; Amin 2011, Eur Respir J; Davies 2012, American Thoracic Society International Conference. San Francisco). Limited longitudinal data from the placebo group of interventional trials are available to define whether an intervention exceeds the intrinsic variability of the test. Obtaining similar data for preschool children and infants is a priority.

#### **16. What is the current evidence? Has computed tomography been reproducibly shown:**

- **to detect even subtle treatment effects ?**

Even using current relatively insensitive scoring techniques, smaller patient numbers are needed to detect treatment effects when chest CT is chosen as outcome parameter compared to routine spirometry (de Jong 2006, Thorax; Tiddens 2007 Proceedings Am Thoracic Soc). Advanced image analysis techniques will increase the ability of CT to pick up small changes.

CT is the gold standard to detect bronchiectasis (Hansell 1998, Radiol Clinics North Am). A single bronchiectatic airway can be easily detected using CT. There are no intervention studies of sufficient duration to show that the development of bronchiectasis can be slowed down or halted. Short term studies over the 2 week course of intravenous antibiotics show that airway wall thickening is reduced and that mucus impaction can be reduced (Davis 2007, AJRCCM). For trapped air small intervention studies showed reduction of trapped air using dornase alpha (Robinson 2005, Chest, Nasr 2001, Ped Pulmonol).

- **if yes over what period of time ?**

Chest CT scoring of bronchiectasis is best suited for longer term studies.

Using CT scoring it has been shown that at the age of 3 years around 30% (Stick 2009, J Pediatr, Mott 2012, Thorax) of children have bronchiectasis, at the age of 5 years this has increased to over 50% (Wainwright 2011, JAMA). The consensus by investigators using chest CT routinely is that to evaluate the effect of a drug on the development of bronchiectasis, a study duration of 2 years is required. More sophisticated image analysis techniques will reduce the number of patients needed in such studies and reduce the study duration to 1 year. Power calculations based on these techniques are on-going and will be completed end 2012

- **that improvements translate into long-term patient benefit?**

Bronchiectasis has been well validated as a clinical relevant outcome measure. It represents an irreversible and well characterised structural change which forms an important component in end stage lung disease. Its presence is a predictor of the number of exacerbation (Brody 2005 AJRCCM, Loeve 2011, Chest) and it has a negative impact on quality of life (Tepper 2010 Ped Pulm). In end stage lung disease its severity is associated with a higher chance of dying on the waiting list for a lung transplant (Loeve 2012, AJRCCM). Prevention of bronchiectasis is thus highly likely to translate to long term benefit. Trapped air is, like bronchiectasis, an important component of end stage lung disease (Loeve 2012, AJRCCM). Its extent is associated with a lower quality of life (Tepper, data on file). How a reduction in trapped air will translate into long term benefit is not as clear as for bronchiectasis.

- **that it can be reproducibly used in multicentre trials across all age ranges?**

All CF centers have CT scanners. SOPs have been developed for pressure controlled CTs under general anaesthesia in young children and are currently used in a multi-centre study in young children in Australia (Clinicaltrials.gov: NCT01270074). In older children spirometer controlled SOPs are available. Funding has been acquired to implement the SOPs by mid-2013 in 12 of the 30 centers participating in the European CF Society Clinical Trial Network. 15 centres have volunteered to participate.

#### **17. Are standardised CT protocols and scoring protocols available?**

For standardised CT protocols, see answer above. The Brody score has been routinely used for over a decade. The CF-CT scoring system is an upgraded Brody scoring system. The CF-CT system is better standardized using a training module, well defined terminology and reference images. The relevance of bronchiectasis and trapped air scores has been well established. The clinical importance of the other component scores is less clear. Using composite total CT score is less likely to be useful. Training sets for the CF-CT scoring system are available. Trapped air on spirometer controlled chest CT can be measured using scoring or with great precision using morphometric methods. More sensitive and accurate (semi)-automated image analysis methods to quantitate bronchiectasis and trapped air are in development and will see the market in 2-3 years.

#### **18. Could CT be used for evaluation of efficacy for treatment of an acute exacerbation?**

In the context of clinical trials chest CT scoring is not the appropriate outcome parameter to study efficacy of treatment of an acute exacerbation. Chest CT has however been used in this context to obtain better understanding of acute events during an exacerbation.

Several studies have shown improvements of chest CTs after treatment of an exacerbation (Shah 1997 AJR; Davis 2007 AJRCCM). However bronchiectasis being the most important irreversible structural change of CF lung disease cannot be reversed. Short term effects on trapped air, airway wall thickness, and mucous impaction can be observed (Robinson 2005, Chest, Nasr 2001, Ped Pulmonol). These observations are useful to improve our understanding of an exacerbation and its treatment, but the clinical significance of these findings for the long term is not clear. In addition, given the radiation exposure the use of short term repeat scans should be restricted. Chest MRI being less sensitive to track morphological changes like bronchiectasis might be of use in the next 5 years to track effects of treatment on mucus impaction.

#### **19. How accurate/robust can an improvement of air trapping predict long-term clinical benefit?**

Trapped air is considered to reflect small airways disease (Tiddens 2010, Pediatr Pulmonol). It is present early in life and it is slowly progressive (Stick 2009, J Pediatrics; Mott 2012 Thorax)). Trapped air needs to be prevented, halted, or reversed. The most important argument for this is that it contributes importantly to end stage lung disease, occupying 10-80% of total lung volume in patients

screened for a lung transplant (Loeve 2012, AJRCCM). In contrast to bronchiectasis, trapped air can be partly reversed (Robinson 2005, Chest, Nasr 2001, Ped Pulmonol) . The extent of trapped air in end stage lung disease is not associated with an increased risk for mortality on the waiting list (Loeve, 2012 AJRCCM).

## **20. What is the cumulative risk of radiation of repeated CT in light of improved longer survival of CF patients?**

This risk has been modelled in several studies and is considered to be low (de Jong 2005, AJRCCM) However, cumulative radiation restricts the use of repeated chest CTs. For clinical routine a chest CT every 2 years is considered low risk. With the new low dose multi-slice CT with iterative reconstruction the radiation dose of 1 CT (0.1-0.2 mS) is comparable to 1 years of background radiation or to the cosmic radiation exposure during 1-2 round trip long-haul flights (Silverman 2008, Lancet) Such low radiation exposure is justified in the context of clinical trials, but still the lowest number of repeat CT's should be aimed for.

## **21. Could a minimal important difference be defined?**

An MCID is not available for this parameter. However, the annual increase of bronchiectasis score has been well defined (de Jong 2006, Mott 2012, Thorax). To reduce the progression of bronchiectasis and or trapped air in a two year study by 30% can be considered clinical significant.

## **22. What is the current evidence of the biomarker "sweat chloride measurement" as potential surrogate endpoint in clinical trials?**

Sweat chloride measurement reflects CFTR function in the sweat gland and is not necessarily a substitute for CFTR function in other organs.

The validity of sweat chloride measurements in diagnosis is well established: it is able to discriminate between patients with CF and non-CF individuals, and between patients with CF and carriers. It is also able to some extent to discriminate between patients with different severity of disease. Clear SOPs are available for semi-automated techniques (pilocarpine iontophoresis for sweat stimulation, sweat collection on the macroduct sweat collection system and colorimetric titration for determination of chloride ions).

However, this outcome measure is only relevant for systemic agents, few of which are in clinical trial to date. The majority of the data is from trials of ivacaftor in patients with the class 3 mutation, G551D:

- In a phase 2 study, this drug led to a dose dependent effect on sweat chloride, reducing it into the borderline/ normal range (Accurso 2010, N Engl J Med).
- In two Phase 3 studies there has been a robust effect paralleled by significant increase in clinically-relevant outcomes (FEV1, exacerbation rate, weight, QoL and LCI) (Ramsey 2011, N Engl J Med; Davies, ECFS 2012). However, in neither trials were there a significant correlation between reduction in sweat chloride and improvement in lung function.
- In the phase 3 trial of PTC124 in patients with class 1 mutations, no effect was seen in sweat chloride and the primary outcome (change in FEV1) was not met (Konstan, ECFS 2012)

## **23. What is the current evidence regarding use of nasal potential difference (PD) or intestinal current measurements (ICM) as potential surrogate endpoint?**

Nasal PD (NPD) and ICM are useful biomarkers of CFTR function in respiratory epithelium and gut in phase II studies but are less useful for phase III studies. Nasal PD is part of the diagnostic algorithm for CF in Europe (De Boeck 2006, Thorax). Mean results of disease groups (CF, non-CF bronchiectasis or controls) and the diagnostic conclusion do not differ with repeated measurements although within-

subject variability is high. SOPs have recently been developed, and equipment is becoming standardized but analysis is not automated. NPD has been widely utilised in clinical trials of both viral and synthetic CFTR gene therapy and has proved the principle of gene transfer and expression (reviewed in Davies & Alton 2010, Proc Am Thorac Soc). Data from studies investigating ataluren, VX770, and VX809 also confirm that NPD is a responsive endpoint in trials that aim to correct the underlying defect in CF. However, as yet, clinical correlates are lacking and the degree of correction in NPD which corresponds to a clinically-relevant outcome is lacking. Data will soon be available from the UK CF Gene Therapy Consortium trial, which may bridge this knowledge gap. Intestinal current measurements are available in specialised centres, and have been used ex vivo to test the pharmacological activity of novel CFTR modulators (for list of refs see De Boeck 2012, Eur Resp J)., but to date, have not been assessed in interventional clinical trials

#### **24. Have other biomarkers been validated and demonstrated to predict patient benefit?**

Many biomarkers have been tested, but information on correlation with clinical outcome or another surrogate outcome been demonstrated for only a few biomarkers (e.g. elastase).

Biomarkers are especially useful in phase II clinical trials because they can provide evidence for the investigational compound's presumed mechanism of action. All biomarkers listed below are not CF disease specific.

A detailed answer regarding benefit would be extremely lengthy and depends on the definition of "benefit"; a) Better Survival; No/possibly (Liou 2012, PLoS ONE). b) correlation to other outcome parameters (many studies, almost in all publications, but low predictive value or specificity). 3. Biomarkers in a) serum, b) sputum, and c) BAL.

##### **Serum:**

Azithromycin reduced circulating neutrophil counts and systemic markers of inflammation, including **C-reactive protein, serum amyloid A** and **calprotectin** in a randomized, placebo controlled trial (Ratjen 2012, Chest.).

##### **Sputum:**

Measurements in sputum have a very high variability even with standardized technique to recover and process sputum. This holds true for cellular components as well as for activities of enzymes and for proteins. Numerous correlations, usually weak and variable have been observed. Recent and classic examples:

**High mobility group box-1 protein** (HMGB-1) predicts incidence and recurrence of acute PE's and survival, (Liou 2012, PLoS ONE).

Initial measurement of **neutrophil elastase** had the highest individual predictive value for subsequent lung function decline while neutrophil elastase, **IL-8** and **IL-6** had the highest combined predictive value. The potential value of elastase in sputum was elaborated by Mayer-Hamblett et al (2007, AJRCCM) by compiling a large database from four multicenter studies: elastase in sputum was negatively correlated with FEV1 ( $r = -0.35$ ; 95% CI  $-0.46, -0.22$ ); on average, patients with CF who differed in their elastase measurements by 0.5 log differed in their FEV(1) values by -7.3% (95% CI: -9.7, -4.6). Lung function decline was associated with increases in neutrophil counts, neutrophil elastase, and IL-1 $\beta$ . (Sagal 2012, AJRCCM in Press doi:10.1164).

##### **BAL:**

**Cellular differential and total neutrophil count** of BAL fluid indicate lower airway inflammation and can be used to assess response to intervention (alpha-1-protease inhibitor) (McElvaney NG Lancet).

1991 Feb 16; 337(8738):392-4) dornase alpha (Paul 2004, Am J Respir Crit Care Med.). Soluble markers have a high variability and do not appear to be very useful; exception may be demonstration of the **presence / absence of active neutrophil elastase** in antiprotease studies (Review in Griese 2008, Eur Respir J).

A useful pulmonary biomarker in the context of CF is measurement of **mucociliary clearance**. There is however need for further standardization of delivered dose and algorithm to calculate clearance.

Measurements of **exercise capacity and fitness**: For an overview of clinimetric properties and feasibility of these biomarkers we refer to Bradley et al. (Bradley 2012, J Cyst Fibros).

### **Questions - Patient related outcomes (PROs):**

#### **25. CFQ-R has been developed in stable CF patients.**

- The CFQ-R is the measure of choice for CF clinical trials; however, it is not advised as the primary outcome parameter.
- The CFQ-R respiratory domain has demonstrated appropriate changes where lung function / exacerbation changes have occurred (e.g. trials of inhaled antibiotics, hypertonic saline, liposomal antibodies, ivacaftor, dry powder colistin, mannitol). The CFQ-R respiratory domain is a symptom score rather than a quality of life score.
- The MCID for the CFQ-R respiratory domain is 4.0 in stable patients and 8.5 in the context of pulmonary exacerbations.
- Which domains of the CFQ-R are appropriate for the study should be decided during study planning.
- There are numerous language translations.
- There are child and adults measures.
  - Limitations
- The CFQ-R was not originally developed for use in clinical trials but to measure the impact of CF disease on QOL.
- The majority (if not all) HRQoL measures used in CF have shown ceiling effects across many domains in patients with mild disease. This is a problem for trials and requires careful patient selection (inclusion / exclusion criteria), or the analysis and interpretation of results needs to take into account the subgroup of patients who would not be expected to respond.
- There are no agreement between childhood data and adult measures. The scales contain different items therefore the data should not be combined in analysis.
- A two week timeframe may not be appropriate.
- Not all language translations have linguistic validity
- Patients do not like long lists of questions; there is a need for a simplified (e.g. 3 questions) global score that measures QOL
- Questionnaires may not be filled out on a daily basis by patients resulting in recall bias when completed before study visits

## **26. Are PRO available or in development to evaluate respiratory symptoms and effects of short treatment courses of acute exacerbations?**

- No such scales are currently available for use as validated outcome measure in clinical trials.

However, there are data that would inform the development of such a PRO.

- Multicentre study in the UK & Ireland using Delphi survey methodology produced a hierarchy of indicators of a pulmonary exacerbation (PE) from both the perspective of adults with CF and CF health professionals. A different hierarchy emerged for the two groups (McCourt et al. in preparation).
- UK Study, using interview methodology, to determine patient-reported indicators of a PE. Child-reported indicators tended to map onto those reported by adults. Patients predominantly focused on respiratory symptoms and how they feel at the onset of a PE but when asked how they recognise when they are improving they focused on what they were now able to do. Patient-reported outcome measures may be more sensitive if they include indicators of 'activity' as well as items of respiratory symptomatology (Abbott 2009, J Cyst Fibros; Abbott 2012, J Cyst Fibros). It is noteworthy that patient-reported *physical functioning* predicted survival (Abbott 2009, AJRCCM). Different indicators of a PE may be associated with different degrees of disease severity and different severity of exacerbation. This requires further exploration.

## **27. What is the current status of the CF Respiratory Symptom Diary? (CFRSD)**

It is currently (Sept 2012) not available for use as an outcome parameter in clinical trials. Evaluation of the measure is on-going. The authors are in the final stages of completing a dossier to go to the FDA, having pooled data from 5 studies with over 400 CF patients.

There are two peer-reviewed published papers:

1. A qualitative paper that describes the generation of the instrument's items. The diary was designed specifically to assess CF symptom severity and frequency. The data are based on 25 interviews (12 adults, 13 children or parents) (Goss J Cyst Fibros 2009;8:245-252). The instrument consists of a 16-item daily diary: 8 symptom items (cough, chest tightness, difficulty breathing, wheeze, coughing up mucus, fever, chills, fatigue) 4 activity impact items (sleep, school / work attendance, reduction of usual activities, time sitting / lying down), 4 emotional impact items (worry, frustration, sad / depressed, cranky).
  2. A 7 day recall of CF symptoms provided similar information to daily diary reporting (Bennett. J Cyst Fibros 2010; 9:419-424).
- Patients do not like completing diaries. The amount of missing data and recall bias may be problematic.

## **28. Are any validated PRO instruments available for pre-school children?**

PRO for preschool children is currently not available.

- A pictorial preschool-version of the CFQ-R (age 3-6 years) is under development and is being psychometrically evaluated.
- It is questionable whether preschool children can report symptom burden and QOL reliably. A ceiling effect is expected to limit the usefulness of PRO in this age category.

## ***Additional issues discussed***

### **Safety endpoints**

Safety assessment should be seen in the context of the entire drug development program. The appropriate drug specific safety parameters should therefore be included in phase III trials. No universal CF specific safety parameter could be defined.

Whenever possible a drug exposure of at least 100 subjects during at least 12 months should be aimed for. To detect rare side effects post marketing, data acquisition in CF patient registries seem to be the most reliable option.

### **Trial designs other than parallel placebo controlled trials.**

Because patient cohorts are limited, there is a need to explore alternative trial designs.

Cross-over design;

Modified N=1 designs;

Adaptive trial design.

Can all be considered, but need careful consideration and discussions with regulatory agencies prior to initiating phase III trials.

The CF databases can be explored to evaluate patient, time and cost savings in alternative trial designs involving modelling. An FP7 program application for this in CF has been submitted.

## **Identified areas in need of further research and education in Group 1**

### ***Education***

Standard schemes for phase 1-3 trials preferred by EMA.

Clearly define necessary characteristics of trials or the different goals e.g. superiority, non-inferiority, etc (e.g. what does non inferiority mean? "not different from placebo" or "is the same as previous verum").

### ***Research***

Information on short- and medium-term variability of outcome measures in stable patients in different age and disease categories.

### ***LCI***

Validate LCI in interventional trials as outcome measure for clinical efficacy; establish link to clinical benefit.

### ***HR-CT:***

Develop more sensitive and accurate (semi)-automated image analysis methods to quantify bronchiectasis and trapped air.

Establish whether reduction in trapped air translates into long term benefit.

**PRO**

Develop and validate PRO instruments for pre-school children.

Develop short, simple, global QOL measures.

**Biomarker**

Define the value of common serum (CRP, IgG) markers in relation to other outcomes in phase 3 trials.

# Annex II of the workshop on endpoints for cystic fibrosis clinical trials

## Breakout group 2: Bronchopulmonary infection

### 1. General microbiological issues:

#### **Background**

- Reliable bacterial detection has inherent problems: sampling technique, oropharyngeal contamination, sputum not homogenous, etc.)
- Recent work using culture-independent methods suggests that CF lungs harbour a vast array of bacteria not conventionally implicated in CF lung disease.
- *In vitro* susceptibility data and clinical outcome in chronically infected CF patients are poorly correlated. [Aaron 2005; Moskowitz 2010; Smith 2003].
- Breakpoints for aerosolized antibacterial agents are unknown.

#### **Questions**

1. **What should be the recommended sampling approach to be used for primary microbiological endpoint?**
2. **What should be the recommended approach for sputum processing? Including:**
  - **time to lab,**
  - **methods for sputum pre-treatment,**
  - **methods for culture,**
  - **assessment of reproducibility and sensitivity.**

#### **Conclusion:**

- Microbiological methodology is important
- Spontaneous or induced sputum is the preferred sample if age allows and limited data suggest that the results with these are generally comparable
- Protocols should set hierarchical schema for alternative methods if no sputa
- There are several national guidelines (UK and Germany) and the Therapeutics Development Network (TDN) of clinical trials in CF in existence regarding best laboratory practice in the CF setting; should be followed and/or use same principles

### 3. **What should be the role of centralised labs?**

#### **Conclusion:**

- Use of centralised labs is preferred provided these are experienced in CF setting but need to establish efficient transportation based on knowledge of effects of delay.

- If local labs are used, there must be lab protocol and appropriate QC system in place.
- 4. **The emergence of other pathogens should be documented in long-term antibacterial agent suppression trials. What diagnostic methods should be used: culture based or culture independent methods? What is the current evidence?**
- 5. **What valuable information can susceptibility testing provide; how to interpret MICs when there is no valid way of setting breakpoints?**
- 6. **Should susceptibility testing be requested as secondary endpoint to evaluate emerging resistance pattern?**
- 7. **In light of poor correlation of In vitro susceptibility data and clinical outcome, should susceptibility testing be included as inclusion criterion for selection of patients?**

#### **Conclusion:**

- Isolation of other species (bacterial and fungal) over time should be sought currently using culture-based methods but this may change in future with on-going improvements in molecular diagnostics; an appropriate definition for treatment emergent pathogens is lacking.
- Not known how many colonies of any one species should be picked up for study, including for susceptibility testing; this needs further evaluation. Multiple picks are recommended.
- There is no scientific basis for setting breakpoints relevant to inhaled antibacterial agents.
- Continue to encourage collection of MIC data and explore relationship with clinical data.
- MICs should not be pre-defined secondary endpoints.
- Patient selection criteria should not include MICs at baseline

## **2. PK/PD studies – particular focus on inhaled antibacterial agents**

### ***Background***

- Pulmonary drug deposition highly variable, attributable to several factors, such as device specifications, the physical properties of the inhaled/nebulised antimicrobial agent, the severity of lung disease, the inhalation technique, breathing pattern.
- Optimal inhaled dose for almost all antimicrobial agents unknown.
- High intra- and inter-subject variability with sputum production and sputum drug concentration, making it difficult to reliably quantify the drug concentration in sputum.
- Observations from clinical trials indicate that increasing MICs do not predict lack of clinical response.
- Cumulative toxicity may become of greater concern, particularly in children, who survive increasingly longer, receiving many courses of antibacterial agents (both inhaled and parenterally).

### ***Questions***

- 8. **What microbiology outcomes are pertinent to the determination of efficacy of aerosolized antibacterial agents?**
- 9. **What is the relationship between sputum drug concentration and efficacy?**

10. How to best define PK/PD profile for inhaled antibacterial agents?
11. How can PK/PD help in defining appropriate dosing for pivotal studies of nebulised antibacterial agents?
12. Are microbiologic endpoints (e.g. cfu/g sputum) useful to determine dosing regimen?
13. Are alternate sputum biomarkers or novel bacteriological tools available?

**Conclusions:**

- Microbiology outcome focuses on cfu/g in respiratory samples.
- There may be some very broad relationship (group level) between sputum density reduction and clinical (lung function) response but it falls down at the individual level.
- PK/PD analyses may look for relationship between dose/PK and various PD endpoints, including cfu/g and FEV1.
- Should also explore PK/PD for safety endpoints if there are relevant safety endpoints to assess.
- Change in cfu/g is often proposed as the main endpoint in dose-finding studies due to practical difficulties in using clinical endpoints to identify potentially useful dose regimens. At the very least the microbiological impact could be used to rule out pursuing products and regimens that appear to have no effect; however, important to stress that the microbiological impact does not predict the clinical response or magnitude of response.
- Should at least document and describe clinical (respiratory) responses in dose-finding studies, even if the study does not plan to base dose selection for Phase 3 on clinical endpoints.

**Questions - Paediatric specific considerations:**

14. What clinical pharmacology data are needed to guide decision whether or not different dosing or regimens among different paediatric age groups might be needed to ensure both efficacy and safety in light of increasing concern of cumulative toxicity?
15. In case of need for lower than adult dose, is weight based dosing appropriate?
16. Tolerability/acceptability considerations, handling of device, frequency of administration impact on the effectiveness of aerosolised antibacterial agents in children. Are standardised PRO instruments/questionnaires available to evaluate those issues?

**Conclusions:**

- Select regimen taking into account systemic absorption issues but may end up with adult dose and regimen depending on the drug.
- Need to define the minimum appropriate age for any delivery system.
- Chronic infection rare/very rare in some age groups which also impacts minimum age.

### **3. Antibacterial agent eradication treatment (AET):**

**Background:**

- AET is now routine in many countries and its success has been documented in several studies.
- Eradication success rates exceeding 90% have been observed.

- The optimal antibacterial agent regimen is not known.
- P. aeruginosa serology has been demonstrated to be a useful marker for successful eradication.
- The incidence of annual P. aeruginosa acquisition is not definitively known; literature suggests a rate between 10% in EU and 15% in the US (Trampers-Stranders, 2010; Treggiari, 2011).

### **Questions - Microbiological endpoint:**

17. Eradication should be the primary endpoint.

**How to define eradication:**

- **With culture-based or culture independent methods ? Have culture independent methods been already validated as outcome measure for use in clinical trials?**
- **What should be the first time point to determine clearance from the airways: one day, one week, other, after end of treatment)?**
- **Should persistence of pseudomonas clearance from the airway be included in primary endpoint? If yes for how long: e.g. 3 negative cultures for 6 months, 12 months, other?**
- **Should Pseudomonas serology be routinely used as secondary endpoint in eradication treatment trials or as a composite/co-primary endpoint?**

18. **What implications for the design of AET studies have recent studies suggesting that CF lungs harbor a vast array of bacteria not conventionally implicated in CF lung disease?**

### **Conclusions:**

- For P. aeruginosa “eradication” should be primary endpoint, based on at least 2 negative cultures starting at least 1-2 weeks post-therapy and at least 2-4 weeks apart.
- Inability to obtain sputa (below a certain age) has implications for certainty re infection status and for demonstrating “eradication” in AET.
- Patients should be followed-up for protocol-defined and justified time period.
- Isolates over time should be typed although not 100% able to differentiate re-growth from re-infection; patient may have started off with > 1 strain and/or may be repeatedly exposed to a reservoir.
- Various typing methods are acceptable such as Variable Number Tandem Repeat (VNTR) and Multilocus sequence typing (MLST).
- Base “eradication” on the initial culture result; accept that this is really “apparent” eradication since negative culture means only that there are insufficient numbers that can be detected by the processing methodology and culture system used.
- No enthusiasm for serology to be routinely used as a secondary endpoint.
- Non-culture methods to detect additional organisms not required but exploration encouraged.

### **Questions - Methodological issues:**

19. **As long as no antibacterial agent drug has obtained marketing authorisation for eradication treatment, how should a clinical trial for regulatory purposes be designed? Is**

**a 28 day placebo controlled trial ethical justifiable and feasible provided all patients on placebo will receive antibacterial agent treatment after 28 days?**

- 20. Is there a need to stratify patients according to their baseline characteristics, e.g. age, respiratory function, microbiology results, etc. ? What would be the most important factors for stratification?**
- 21. In case a comparator will be available, non-inferiority trials against comparator most likely to be requested by regulators; this has implications on trial design and sample size:**
- 22. Could other than non-inferiority trial designs be envisaged, e.g. a non-comparative trial requiring a certain predefined threshold value to be met, e.g. 90% eradication rate?**
- 23. If yes:**
  - **What threshold to request for any new antibacterial agent eradication therapy?**
  - **What could be the biases deriving from the patient population included?**
- 24. The optimal duration of AET is not known. With current evidence, what treatment duration can be considered standard of care and should be requested for regulatory trials?**
- 25. What additional endpoints acceptable for HTA should be included as secondary/additional endpoints in regulatory trials to assess real life benefit?**

#### **Conclusions:**

- Accepted problem of using placebo in light of clinical practise – reflects premise that AET has clinical value since it delays onset of chronic infection; therefore an active comparator is expected to be the only acceptable design for investigators and patients.
- Lack of evidence of the magnitude of the treatment effect to set rational NI margin.
- An adequately powered NI study would require large numbers and may not now be feasible based on current experience of enrolling into AET studies.
- Not clear what would be an acceptable alternative study design. Due to enrolment difficulties maybe a non-comparative study could be further discussed with a pre-defined acceptable threshold and estimate of precision (acceptable criteria need to be adequately justified).
- Stratification should be avoided or minimal - anyway small numbers/cell. However, efficacy so far has mostly been determined in children; reassurance warranted that AET regimens are effective in adults.
- Not clear what duration of treatment should be; probably not less than 1 month.
- Need to establish a definition for an acutely infected CF patient i.e. define a minimum period between routine clinic visits prior to study entry to provide evidence that time elapsed since detection of a new species can be defined in the patient selection criteria (e.g. within prior 2 months could be suitable).
- Currently, the active comparator (if there is one) should probably be 1 month inhaled tobramycin or Copenhagen ciprofloxacin/ colistimethate sodium regimen.
- Aside from the primary microbiological endpoint, not entirely clear how such AET trials could generate data on other endpoints, including endpoints of most interest to HTAs.

- Difficult to define progression to chronic infection – how many courses and/or how many attempts at AET that fail constitute progression to chronic infection?

## 4. Chronic suppressive antibacterial agent treatment:

### **Background**

- Lung damage secondary to chronic infection is the main cause of death. Maintenance of the CF patient's lung function over extended periods of time in the presence of the persistent pathogen in the airways is considered key in the antibacterial agent management of chronic lung infection.
- Antibacterial agents should demonstrate that they achieve the primary goal of treatment of the pulmonary disease: to improve/maintain respiratory function.
- Majority of chronically infected patients are no longer treatment naïve (as they were in the initial TOBI trial), therefore the same effect size may not be expected; this actually was the case in the Colobreath/TOBI comparative trial.  
([http://www.ema.europa.eu/docs/en\\_GB/document\\_library/EPAR\\_-\\_Public\\_assessment\\_report/human/001225/WC500123693.pdf](http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Public_assessment_report/human/001225/WC500123693.pdf)).
- Prior exposure and response to an antibacterial agent which is then used as a comparator may influence the interpretation of the results.
- Improved symptomatic treatment options resulted in preservation of good lung function over prolonged period of life with the majority of (young) children having spirometry within the normal range, rarely experiencing pulmonary exacerbations and a slower yearly decline of FEV1.
- Prevalence of chronic pseudomonas infection lower than previously published: (Rosenfeld 2001 reported prevalence of *P. aeruginosa* of 18% at 1 year and 33% at 3 years of age; Wainwright 2011: 12 % at age 5).
- The rate of pulmonary exacerbations requiring i.v. therapy is declining and lower than previously published (Treggiari et al. 2011).
- Lack of universally agreed definition of acute exacerbation.
- Lack of (universally agreed) standardised criteria when to start parenteral antibacterial therapy; thus initiation of parenteral treatment is not a reliable indicator of exacerbation.
- Several inhaled antibacterial agents are currently authorised for treatment of chronic *Pseudomonas* infection; thus long-term placebo-controlled trials cannot be performed.
- Patient pool available for clinical trials limited; the need for large clinical trials of one product may hinder enrolment of patients into other trials, thus affecting opportunities to study another product.

### **Questions - Outcome measures / microbiological issues**

#### **26. What is the current yearly decline in FEV1 in patients not yet chronically infected with *Pseudomonas* versus those chronically infected?**

#### **Conclusions:**

- Yearly decline in FEV1 is maybe 1-2% but varies by age group; not likely to see anything other than stabilisation in older age groups.

27. **The results of recently approved inhaled antibacterial agents (Cayston, TOBI Podhaler) seem to support the use of FEV1 as primary endpoint in placebo-controlled trials. But is spirometry (FEV1) still sensitive enough to demonstrate treatment effect in patients stabilised on chronic suppressive inhaled antibacterial treatment in non-inferiority trials against active comparator?**
28. **Minimal important differences for respiratory domain in CFQ-R have been defined (Quittner 2009). Would a PRO be acceptable for regulators as primary endpoint if supported by lung function data and clinical data?**
29. **What additional endpoints acceptable for HTA should be included as secondary/additional endpoints in regulatory trials to assess real life benefit?**

**Conclusions:**

- Change in FEV1 from baseline should remain primary endpoint; CFQ-R should show a consistent trend vs. FEV1.
- Insufficient data (and access) at present to supplant FEV1 with lung clearance index (LCI).
- Secondary endpoints may include LCI, time to documented and pre-defined exacerbations, respiratory related hospitalisation, use of IV antibacterial treatment (oral more dubious), measures of burden/convenience.
- If FEV1 is primary endpoint after single cycle (measured on day 28) this will not currently satisfy all regulators (e.g. some may want multiple cycle data on time to exacerbation and FEV1 even though cycling only with the test agent not considered feasible now).

**Questions - Methodological issues:**

30. **A universally agreed definition of chronic Pseudomonas infection would help make trial results comparable.**
- **Can the Leeds criteria (Lee et al, 2003) be universally agreed for defining chronic Pseudomonas infection for regulatory trials?**
  - **Are any other definitions available?**

**Conclusions:**

- A universal definition of chronic infection would be desirable but cannot simply apply Leeds or Copenhagen definitions because clinical practise re culture intervals variable.

31. **In light of poor correlation of In vitro susceptibility data and clinical outcome, should Pseudomonas susceptibility testing be included as inclusion criterion for selection of patients?**

32. **What would be the most important factors for stratification? Which baseline factors are known to affect mostly outcome?**

**Conclusions - Study population:**

- Susceptibility test results should not be part of the inclusion criteria.
- Reasonable to stratify by age and FEV1; other factors up to sponsor.
- May consider enriching population and enhance likelihood of detecting change in FEV1 according to number of past exacerbations.

- Definition of exacerbation and reliability of pre-study documentation questionable.
- 33. Are non-inferiority trials against comparator, requiring larger sample sizes, the only reliable study design?**
- 34. What non-inferiority margin is acceptable?**
- 35. Could other than non-inferiority trial designs be envisaged, e.g. a non-comparative trial requiring a certain predefined threshold value to be met, e.g. a predefined improvement of lung function?**

#### Conclusions:

- Not possible to pre-define a NI margin that is unquestionably robust.
  - Compromise between evidence and feasibility.
  - NI margin should take into account the population characteristics.
  - Can use a pre-planned hierarchical testing (if meets NI then test for superiority).
  - Probably no good alternative to some sort of NI study design but leave options open for the future.
- 36. Inhaled antibacterial agents currently authorised and thus candidates for active comparators were studied in CF population with baseline FEV1 % predicted between 25% and 75%. Particularly in paediatrics this population no longer representative as majority of patients has lung function above 75% pred.**
- **Would regulators accept to also include patients with FEV1 >75% in both new and comparator arm?**
  - **If no: what are the options?**
- 37. Cycled 28 days on/off drug administration in TOBI trial, despite somewhat arbitrary, set precedent:**
- **Can this administration schedule still be considered to be the optimal one?**
  - **Should novel inhaled antimicrobials also use this on/off regimen?**

#### Conclusions - Study design:

- Do not set upper FEV1 limit
- Prospective single cycle studies vs. placebo not at all likely to be possible
- Consensus that studies of 3 cycles over 6 months with randomisation to test agent or single comparator may not be feasible (studies ongoing/recently completed probably last that can use this design) or at least only in some countries because of the current practise of cycling different treatment modalities over time
- One approach might be to use alternative to TOBI for a few cycles and then randomise TOBI vs. test agent but patients may not be willing to enrol into such studies
- If the primary analysis is at 28 days how can we obtain longer-term data on safety and efficacy if routine practise is to cycle the antibacterial agents? May not be able to get pristine agent-specific data for > 28 days
- On-off cycling not proven to be optimal and many centres are using continuous antibacterial therapy but with cycling of the agent (i.e. switch every 28 days)

- For new agents, new formulations and new regimens the only feasible and rational comparator may be continuous and cycled agents as per a well justified standard of care.
- Some such justification may come from registries; academic community could make major contribution in this respect
- Over time describe impact on selection of organisms with high MICs and specific mechanisms of resistance

**38. Novel devices for nebulised antibacterial agents, e.g. e-flow, are widely used in clinical practice but are not authorised. Comparators may only be used with the device authorised (e.g. for TOBI inhalation solution the LC plus inhaler).**

- **Are clinical trials still feasible (patients still willing to enrol) when the use of "old", less convenient, but authorised devices is required for regulatory trials?**

**Conclusions:**

- May be difficult to demand that comparative agents are used ONLY with the device with which they were studied and hence approved.
- Nevertheless sponsors may still feel obliged to use only the licensed devices in pivotal trials.
- Limitations of well documented evidence re use of licensed agents with alternative devices.

**39. In light of the pipeline with several new antibacterial agents about to be tested in CF, secondary endpoints addressing potential added benefits, e.g. reducing the treatment burden, should be added into regulatory trials:**

- **Does the available CFQ-R adequately assess impact on treatment burden?**
- **What other factors could be of added benefit?**

**Conclusions:**

- Secondary endpoints could include documentation of administration time per individual dose and overall per day.

## Identified areas in need of further research and education in Group 2

### ***Number of picks:***

Not known how many colonies of any one species should be picked up for clinical studies, including for susceptibility testing; this needs further evaluation.

### ***Relationship MIC – clinical outcome:***

Collection of MIC data and to explore relationship with clinical data strongly encouraged

### ***Definition of progression to chronic infection:***

How many courses and/or how many attempts with antibiotic eradication treatment that fail constitute progression to chronic infection?

### ***Definition for an acutely infected CF patient***

Need to establish a definition for an acutely infected CF patient i.e. to define a minimum period between routine clinic visits prior to study entry to provide evidence that time elapsed since detection of a new species can be defined in the patient selection criteria (e.g. within prior 2 months)

# Annex III of the workshop on endpoints for cystic fibrosis clinical trials

## Breakout group 3: Exocrine pancreatic insufficiency

The group discussion focused on two issues:

- New pancreatic enzyme products;
- CFTR modifiers and their impact on GI outcomes/endpoints.

### 1. Exocrine pancreatic insufficiency

#### **Background**

- Exocrine pancreatic dysfunction affects approximately 90% of patients; the deficiency manifests as fat and protein indigestion.
- Morbidity and mortality are related to nutritional status.
- Although pancreatic enzyme replacement therapy (PERT) significantly reduces fat and nitrogen in stools, it does not completely correct the abnormal fat absorption.
- With the availability of new genetically engineered single pancreatic enzyme preparations the contribution and digestion with PERT of all three macronutrients to the nutritional status has to be justified.

#### **Questions**

1. **How should lipase products be investigated: is equal potency for fat absorption sufficient or should protein and carbohydrate absorption be also investigated?**

#### **Discussion:**

Dependent on the product: For porcine derived PEPs with defined lipase, protease and amylase levels, there is ample evidence that fat absorption measurements should be sufficient as a surrogate. For potential future recombinant enzyme preparations, which are likely to contain bacterial enzymes and a very limited selection of the wide spectrum of the enzymes secreted by the pancreas Measurements of protein absorption, and carbohydrate absorption are minimum requirements and methods to look at other nutrients should be considered if methods exist that do not require multiple sampling.

#### **Conclusion:**

- Formally, all three components of pancreatic digestive enzymes are needed, especially the protein component (Coefficient of nitrogen absorption).
- If developed as full substitution demonstration of efficacy and safety as regards protein and carbohydrate (in addition to fat-) absorption is essential.
- Lipase only products can only be an add-on treatment (thus a different indication is regarded to be suitable for these products: improvement or normalisation of fat absorption).

- The problem of fibrosing colonopathy with high doses will most likely need to be addressed in post-marketing safety studies (PMS-studies).
- Carbohydrates: Uncertainty regarding the need, but absorption related to enzyme preparations has been tested. Future research has to be awaited.

## **2. For new pancreatic enzyme replacement therapy (PERT) formulations: is faecal fat the investigation of choice or can breath tests be alternatives?**

### **Discussion:**

Breath tests use only one substrate in contrast to fecal fat presenting the sum of enzyme activity on a variety of substrates.

### **Conclusion:**

- Currently, faecal fat (CFA) has to be regarded to be gold standard and validation data on breath tests are insufficient to recommend these to replace CFA.
- Alternatives to CFA, however, should be further investigated and evaluated due to the well-known set-backs of the CFA. Qualification advice is recommended.
- For breath tests future research should evaluate the influence of pulmonary function (lung disease severity).

## **3. Coefficient of fat absorption (CFA), considered being the gold standard for the diagnosis of fat malabsorption, is not sensitive enough for detecting only small treatment effects; companies therefore have declined studies for this reason in submitted paediatric investigation plans. The PDCO considers requesting comparison of CFA to C13 FA and or C13 TG tests. At present hydrogen breath tests are not reliable enough; but is this because of too small number of studies conducted so far or because hydrogen breath test really is not able to reliably detect small changes?**

### **Conclusion:**

- Both reasons mentioned are considered applicable. Due to the small number of studies and the different scope of these studies, it is currently unclear whether the tests will be able to detect small changes. The (potentially marginal) clinical relevance of small differences/changes has, however, to be kept in mind.

## **4. Which test can serve for dose finding in PK/PD studies?**

### **Discussion:**

To date, EPI dose is titrated individually by symptoms. Furthermore, enzymes act locally. Therefore, PK/PD studies do not make much sense because clinically patients are titrated on clinical symptoms, but not on acid steatocrit. Therefore, the "dirty" method of titrating by symptoms cannot quite be abandoned.

### **Conclusion:**

- PK studies probably do not make sense at all.
- PD: Acid steatocrit and breath tests could be a method for dose-response (recommended for new products).
- Symptom score/stool consistency to define dose ranges?
- Dose-finding not very sensitive.

- Improvement of persistent steatorrhoea after optimal PERT with control of clinical symptoms might be considered in add-on therapy, because of different mechanism of action.

**5. Is it necessary to test children below the age of 6 years or can adults be tested and can adult results be extrapolated to children?**

**Discussion:**

Adult results can be extrapolated to children since dose titration needs to happen individually, anyhow, for standard porcine-derived PEPs. For future recombinant enzyme products, little will be known about the desired and necessary effect on growth and nutritional status, in this class pediatric studies with a clinical endpoint of growth are necessary.

**Conclusion:**

- For “me too” PERTs, the inclusion of children is not needed.
- For new (recombinant) products, separate studies in all age groups are needed.
- Optimisation of dietary measures like increase of protein content should be implemented, when improvement of growth will be evaluated.

**6. How to best control for differences in food intake as variable influencing the test results of faecal fat measurements?**

**Discussion:**

This is a difficult problem. Even if fat intake is standardized, the method of calculating fat intake varies widely across clinical sites, so that a central dietician is needed, which in turn creates problems for sites where the local method to calculate fat intake gives very different results from the central method with regards to the clinical practice of that site outside the study. Also, changing the diet of an individual upsets the PEP regimen that patient used prior to entry into a study, necessitating new dose titration.

**Conclusion:**

- Food intake is important: should be controlled, standardized, and documented.
- Standardisation for children difficult, but possible (different meals with similar fat content).
- Problem: Standardisation within dieticians.

**7. The current CF guideline requests placebo-controlled superiority trials in the frame of add-on studies on top of standard therapy. The primary endpoint must be clinical. Are superiority trials compared to standard of care (currently available enteric coated porcine preparations) realistic and feasible? Are non-inferiority trials in light of limited patient population feasible?**

**Discussion:**

Placebo-controlled studies are – in a “me too situation” unethical and practically not feasible. Given the high efficacy of current porcine-derived PEPs a superiority trial will likely require similar sample sizes as non-inferiority trials, unless a highly selected population is included.

**Conclusions:**

- Add-on: only in the setting of add-on lipase, or protease. Could be achieved by looking at CFA first and then prove that fully normalised CFA is clinically superior. Superiority should be possible to be achieved when patients are selected carefully.

- New “full” PERTs: no placebo control possible.
- Consequence is a clinical equivalence or non-inferiority trial for which currently the non-inferiority margin, and consequently the sample size is unknown.
- Alternative could be baseline CFA without PERT in all participants followed by chosen treatments.

## 8. Is there a need for age appropriate formulations for infants?

### Discussion:

Experience is as long as caregivers do not need to “tamper” with dosing (e.g. measure dose with a spoon or split capsule content), the current method to sprinkle beads from one low strength capsule onto acidic food is adequate. A bigger problem is, older children (toddlers) as they start to need higher doses, i.e. many capsules or content from many opened capsules. There, intake of 10 ml liquid vs content of e.g. 10 capsules would be desirable.

### Conclusion:

- Yes, liquid formulations are clearly needed.

## 2. CFTR modifiers - impact on GI outcomes:

### Background

- For the first time CFTR modifiers are available and are assumed to improve CFTR dependent absorption mechanisms.
- Weight gain (as observed in ivacaftor trial) not necessarily results from improved absorption.
- Can improvement of CFTR functions in the GI tract be quantified? Is it possible to define meaningful digestive endpoint (s)? If yes, could it/they be used as surrogate for clinical efficacy in preschool children in whom repeat lung function tests are not (yet) feasible?
- The greatest expected benefit of such therapy would be expected in young children as the translation of disease improvement into improved organ function may be limited by the level of irreversible damage at the time of treatment initiation.

### Questions

9. **When evaluating CFTR modifiers, which of the following GI functions/tests should be assessed: faecal fat or fatty acid stable isotope absorption, bile acid re-absorption, protein absorption, breath tests? Or is a combination more appropriate?**

### Discussion:

In initial studies, a combination is more appropriate since it cannot be assumed that the level of CFTR modulation has the same effectiveness in all organs to correct the function. So at least CFA, CNA and bile acid reabsorption should be studied.

### Conclusion:

General:

- Fat, protein-, carbohydrate absorption, should be evaluated.

The following additional investigations may be of use:

- rectal biopsy: intestinal electric currents (“Ussing chamber”)
- “organoids”: intestinal crypt culture system (F-508del mutant mimicking)

Regarding different test on organ function the following is desirable:

- Liver: (Cirrhosis in 10% only)
  - Ultrasound, Fibroscan, liver biopsy.
- Intestine:
  - Fat excretion (quantitative stool tests, fatty acid absorption, protein, amino acid, carbohydrate absorption (stable isotope absorption tests, stable isotope breath tests);
  - Permeability (51CrEDTA, Lactulose/Mannitol absorption test), although relevance unclear at present;
  - Intestinal pH (wireless motility capsule);
  - Intestinal Passage time (also: Smart Pill); DIOS incidence;
  - Bacterial overgrowth with adequate test system (including microbiota);
  - Intestinal inflammation (calprotectin, endoscopy, biopsy, capsule endoscopy; rectal NO production);
  - Bile acid malabsorption (unclear relation to CFTR malfunction; relates to CFTR/ASBT interaction): Synthesis measures (isotope dilution), FGF19, C4.
- Pancreas :
  - Faecal elastase (in newborns for potential “rescue”);
  - Intestinal pH (wireless motility capsule);
  - Intestinal Passage time (also: Smart Pill); DIOS incidence.
- The extent of necessary investigations depends on the indication(s) aimed at, and potential additional claims aimed at by the Sponsor/Pharmaceutical Company. The choice of indication, and additional claims, however, is at the discretion of the Pharmaceutical Company. A selected range of extra pulmonary investigations, however, is desirable, even if only a pulmonary indication without additional claims is aimed at. Only collaboration with gastroenterological experts and the sponsor will enable validation of a panel of multiple tests that represent in the end a meaningful contribution to CFTR modifier evaluation.

#### **10. Should pancreatic function be tested and how: by elastase, chymotrypsine ?**

##### **Discussion:**

Stool elastase is a very unreliable method with little predictivity of EPI severity. Chymotrypsine might not be predictive for restoring of the activity of multiple enzyme classes. Ideally pancreatic function should be studied by a secretin test.

##### **Conclusion:**

Not chymotrypsine, elastase possible; bicarbonate secretion by pH in duodenum. An increase in pH could be assessed by capsule endoscopy.

#### **11. Should liver function be tested and how: can improved biliary bicarbonate secretion be tested?**

**Discussion:**

Clinically, bilirubin and LFT should suffice. It could be explored whether biliary bicarbonate and enzyme secretion, resulting in an increase in pH, could be assessed by capsule endoscopy.

**Conclusion:**

- Routine lab sufficient (ALT, AST, bili).

**12. Can one of the tests for evaluation of maldigestion serve as a new surrogate for clinical benefit of CFTR modifiers in children below the age of 6 years, who are unable to perform FEV1 measurements?****Discussion:**

It is not expected that maldigestion endpoints can be indicative of lung function improvement. It cannot be assumed that the level of CFTR modulation has the same effectiveness for all organs to correct the function.

**Conclusion:**

- No, no data on organ specificity known.

**13. Which test can serve for dose finding in PK/PD studies?****Discussion:**

In case of dose finding for CFTR modifiers for correcting maldigestion: If biliary bicarbonate secretion, measured as pH increase, could be established as a surrogate endpoint also for pancreatic function, it could be a sensitive method for establishing PK/PD relationships. Otherwise, to validate acid steatocrit as a surrogate endpoint.

**Conclusion:**

- Currently unclear, because only partial validation data are available.

**14. Should/could chloride sweat measurements and/or nasal PD be used as reference to validate those tests?****Discussion:**

The effect of CFTR correction/modulation in different end organs is currently unknown.

**Conclusion:**

- Currently not recommended; when compared, organ system differences can be assessed.

**15. Is it necessary to test children below the age of 6 years or can adults be tested and can adult results be extrapolated to children?****Discussion:**

Children need to be tested given the reasons mentioned in the background section (irreversible loss of function as disease progresses, hence there is a low predictive value of drug effect in patients with progressive loss of function).

**Conclusion:**

- No, extrapolation is not possible, sequential testing with small children in later stages necessary. Neonatal screening might result in a need to test younger children already at early drug

development stage because of potentially higher preservation of organ function in patients diagnosed through neonatal screening.

**16. Should infants be studied separately, since they still might have some potential for exocrine pancreatic function reversibility?**

**Conclusion:**

Yes. (Also refer to answer to Q15)

- 

**17. The degree of CFTR insertion in the epithelial cell membrane induced by CFTR modifiers is assumed to vary strongly between different CFTR gene mutations, as has been observed with ivacaftor. How can efficacy be evaluated when only small amount of CFTR insertion is expected? How can responders be differentiated from non-responders?**

**Discussion:**

As already discussed for Q 13: surrogate parameters should be established. These are mostly investigations in the early phases of drug development. Depending on the outcome (e.g. if the in-vitro effect is sufficiently strong), a choice of selected organ functions can be tested at later stages of development. The potentially very small subgroups do, however, restrict the use of efficacy evaluations at later stages.

**Conclusion:**

- Inclusion of the GI tests questionable, but if potential groups are big enough may be yes. Also it might be necessary because of differences within organ systems.

**18. The expected improvement of bile acid reabsorption through CFTR modifiers might influence the already enlarged BA bile acid pool (through compensatory over production by the liver). Should this be taken into account in the design of clinical trials by also measuring bile acid pool size and decrease due to effects of CFTR modifiers?**

**Discussion:**

Generally yes, but predominantly for safety reasons.

**Conclusion:**

- Effects on bile acid pool can currently only be assumed. Therefore, the potential for an influence on bile acid pool should be further investigated.

**19. Should the influence of luminal alkalization been studied as well?**

**Conclusion:**

- Yes, it is currently not known whether it is a decisive factor.

**20. Is there a need to evaluate whether CFTR treatment has an impact on occurrence of DIOS, diabetes?**

**Conclusion:**

- DIOS yes, since improved CFTR mediated chloride secretion might have a clinical effect.
- Diabetes yes.

## Identified areas in need of further research and education in Group 3

- Use of breath tests in CF patients: the influence of pulmonary function on breath test results should be evaluated.
- Need to validate acid steatocrit as a surrogate endpoint.
- Further validation of electrophysiological in vitro tests for CFTR modulation.
- Further validation of clinical tests for liver, intestinal, and pancreatic organ function in the condition CF.
- Effect of CFTR modulators on bile acid pool size.