

Outcome Measures for Cystic Fibrosis

Gene Therapy Clinical Trials

A thesis submitted for the degree of Doctor of Philosophy

December 2012

Dr Gwyneth Davies

MBChB MSc(Res) MRCPCH

Department of Gene Therapy

National Heart and Lung Institute

Imperial College London

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence.

Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work

Declaration of Originality

I declare that the work within this thesis is my own unless stated otherwise and that external sources are acknowledged appropriately.

The work of this PhD has taken place within the UK Cystic Fibrosis Gene Therapy Consortium's (UKCFGTC) clinical programme. This declaration page therefore includes an explanation of the role of Consortium members from the three UK centres (Imperial College London, University of Oxford and University of Edinburgh) in the two clinical studies included in this thesis (the Run-In study and the single dose CF gene therapy Pilot study). Additional description and detail is also provided within individual chapters.

Run-In study: Recruitment to the Run-In study was complete prior to the author joining the clinical programme; however the study has been in progress throughout the time course of this PhD. Over this period, the author's role has been predominantly in overseeing and refining the methodology of airway physiology tests, quality control and data analysis, widely implemented across the Consortium throughout the time course of this PhD. The clinical team at each site included clinicians and research nurses at the University of Edinburgh, and clinicians, nurses and physiologists at the London site. Scientists and administrative staff were based at all three Consortium sites (the Oxford site is non-clinical).

Pilot study: The clinical team at the London site was led by Dr Jane Davies (Reader, Honorary Consultant at the Royal Brompton Hospital, London and primary supervisor for this PhD). Other members include nurses, physiologists, trial co-ordinator and clinical research fellows (including the author of this thesis). The scientists at the London site were led Dr Uta Griesenbach, and the Chief Investigator was Professor Eric Alton. Sample processing (sputum and serum) was performed by laboratory technicians within the Department of Gene Therapy, Imperial College and at the Edinburgh Consortium site. Routine clinical laboratory tests (blood and sputum) were processed by the Royal Brompton Hospital laboratory services. Molecular assays of *CFTR* were performed at the Oxford site of the Consortium, by the team of scientists led by Dr Steve Hyde. The Pilot study took place throughout the first 2 ½ years of the author's PhD. The author had pivotal roles in the conduct of the trial: consenting patients, providing medical care for patients on bronchoscopy and dosing days, and performing airway potential difference (PD) measurements. During the study, the author also prepared summary progress reports for the UKCFGTC Strategy Group, Data Safety Monitoring Board (DSMB) and regulatory authorities. Post trial, the author's focus has been on data analysis. This has included all upper and lower airway PD measurements, as well as safety and airway physiology outcomes.

Abstract

Background: Cystic fibrosis (CF) is a life-shortening, chronic respiratory disease caused by mutations in the *CFTR* gene. Novel therapeutic agents such as gene therapy aim to correct *CFTR* and to demonstrate evidence of molecular, functional and (ultimately) clinical efficacy. It was hypothesised that currently used methods to detect these changes may be optimised to enhance sensitivity and allow quantification, and facilitate an understanding of geographical effects within the airway.

Methods: Outcome measures were investigated within two UK CF Gene Therapy Consortium studies; a longitudinal observational study ('Run-In'), and a single dose gene therapy study ('Pilot') which investigated safety and functional efficacy. Measurement of airway function with spirometry and lung clearance index in the Run-In study allowed investigation of variability and change over time and comparison with other outcomes. Development of methodology and data analysis from measurements of potential difference (PD) in the nose and lung in the Pilot study allowed investigation of these as measures of functional efficacy.

Results: In the Run-In study, the choice of external reference source was crucial for interpretation of spirometry outcomes. Airway physiology outcomes correlated with structural changes on chest CT however were limited in their ability to detect site of airway abnormality. There was some evidence that disease severity was associated with intra-subject variability and affected rate of change over time. In the Pilot study, airway PD was shown to change post gene therapy within individuals but the responses were not universal and depended on the definitions used. A novel method of nasal PD quantification did not improve an ability to quantify change.

Conclusions: There is no single universal outcome measure in CF, but it is important to take account of the patient population in terms of disease severity. Whilst it would be inappropriate to relate PD outcomes with clinical outcomes in the Pilot study; this will be an important relationship to understand in the future in order to allow rational design of CF gene therapy clinical trials.

Acknowledgements

The opportunity to be part of the UK Cystic Fibrosis Gene Therapy Consortium's clinical programme has, in my opinion, been an unrivalled experience. The collaboration between scientists, clinicians and administrative staff at all sites is a real example of effective teamwork and one to which I am proud to have played my part. In particular, the Clinical Trial co-ordinators at the London site (Nia Voase and Clare Saunders during the time course of this PhD) deserve thanks.

Clinical trials are not possible without the commitment of patients who give their time in an altruistic manner for the purposes of research, I thank them too.

Throughout my PhD I have had the support of my four PhD supervisors - Professor Eric Alton, Dr Jane Davies, Professor Andy Bush and Dr Mark Rosenthal. Each brought something different to my perspective and approach to research that will stay with me for many years. My particular thanks goes to Jane, who has been steadfast in her support and encouragement over the past 4 years and put simply, is just absolutely inspirational.

Table of Contents

1	<i>Defining the context</i>	20
1.1	Clinical evaluation of novel therapeutic agents	20
1.2	Cystic fibrosis	20
1.2.1	CFTR	21
1.2.2	Morbidity and current approach to management	22
1.3	Gene therapy for cystic fibrosis	22
1.3.1	Vectors	24
1.3.2	Clinical Trials	28
1.3.3	Challenges	31
1.4	Why we need ‘better’ outcome measures for CF clinical trials	32
1.5	The UK CF Gene Therapy Consortium clinical programme	34
1.5.1	The Run-In study	35
1.5.2	The single dose ‘Pilot’ study	36
1.6	Outcome measures of interest	37
1.6.1	Airway physiology outcomes	37
1.6.2	Airway potential difference (PD)	41
1.7	Hypotheses, aims and objectives	45
1.7.1	Hypotheses	45
1.7.2	Aims	45
1.7.3	Objectives	46
2	<i>Methodology</i>	47
2.1	The Run-In study	47
2.1.1	Study personnel	47
2.1.2	Study recruitment and consent	47
2.1.3	Study visits and clinical stability	49
2.1.4	Study protocol	49
2.1.5	Data analysis	50
2.2	The single dose Pilot study	52
2.2.1	CFTR gene therapy product pGM169/GL67A	52
2.2.2	Trial protocol	53
2.2.3	Study enrolment	54
2.2.4	Ethics, MHRA and study protocol amendments	56
2.2.5	Study visits	58
2.2.6	Dosing: Preparation	59
2.2.7	Dosing: Gene therapy administration	61
2.2.8	Clinical monitoring	63
2.2.9	Adjunctive therapy	63
2.2.10	Adverse events and safety reporting	65
2.3	Outcome measures: Airway physiology	65
2.3.1	Spirometry	65
2.3.2	Lung clearance index	75
2.3.3	Gas transfer	83
2.4	Outcome measures: Clinical, laboratory, structural and patient reported	84
2.4.1	Physical examination and pulse oximetry	84
2.4.2	Clinical laboratory blood tests and genotype	85
2.4.3	Structural: Chest computed tomography (CT) scan	85
2.4.4	Patient reported: Quality of life questionnaire	88
2.5	Functional efficacy outcome measures: Potential difference measurements in the upper airway	90
2.5.1	Measurement of nasal PD: overview	91
2.5.2	Patient testing and detailed description of set-up	94
2.5.3	Methods: data analysis	101
2.5.4	Statistics	106

2.5.5	Patient inclusion in nasal dosing cohort.....	106
2.5.6	Novel approach to nasal PD testing.....	107
2.6	Functional efficacy outcome measures: Potential difference measurements in the lower airway	110
2.6.1	Previous methodology and rationale for changing existing set up	110
2.6.2	Methods and initial early phase results	112
2.6.3	Evaluation of the novel set up and use in Pilot study	117
2.6.4	Further development and use in Pilot study	119
2.6.5	Data recording and analysis in Pilot study	121
2.6.6	Statistics	124
2.7	Functional efficacy outcome measures: molecular evidence of <i>CFTR</i> gene transfer.....	125
2.7.1	Nasal brushing for molecular CFTR assays	125
2.7.2	Bronchial brushing for molecular CFTR assays	125
3	<i>Results: Run-In study</i>.....	127
3.1	Patient recruitment and characteristics of study population.....	127
3.1.1	Subject characteristics.....	128
3.1.2	Interval between visits	129
3.2	Description of data	130
3.3	Summary data and variability	132
3.3.1	Physiology: Spirometric values.....	132
3.3.2	Physiology: Lung clearance index	135
3.3.3	Structure.....	136
3.3.4	Inflammation.....	137
3.3.5	Quality of life questionnaires	138
3.3.6	Summary	140
3.4	Cross-sectional correlations within categories	141
3.4.1	Airway physiology	141
3.4.2	Structure.....	144
3.4.3	Inflammation.....	146
3.4.4	Quality of life	146
3.5	Cross-sectional correlations between categories	146
3.5.1	Physiology vs. structure	146
3.5.2	Physiology vs. inflammation.....	149
3.5.3	Physiology vs. Quality of life	149
3.5.4	Structure vs. inflammation.....	149
3.5.5	Structure vs. Quality of life	149
3.5.6	Inflammation vs. Quality of life	151
3.6	Longitudinal analysis: change over time	151
3.6.1	Physiology	151
3.6.2	Inflammation.....	154
3.6.3	CFQ-R Respiratory domain.....	155
3.6.4	Influences on longitudinal change.....	155
3.6.5	Correlation of longitudinal change between outcome categories.....	157
3.6.6	Change over time with data from < 4 study visits	158
3.7	Discussion	159
3.7.1	Airway physiology analysis methodology.....	160
3.7.2	Recruitment and retention	161
3.7.3	Defining clinical stability	162
3.7.4	Data availability	162
3.7.5	Coefficient of variation	163
3.7.6	Ceiling effect	164
3.7.7	Change in outcome measures over time	164
3.7.8	Relationship between structural change on CT and other outcomes	166
3.7.9	Patient reported outcomes.....	167
3.7.10	Variability according to disease severity	167
3.7.11	Ability to detect change	167
3.7.12	Moving forward to multidose trial	168

4	<i>Results: Single dose Pilot study</i>	169
4.1	Recruitment	169
4.2	Baseline characteristics	169
4.2.1	Lung gene therapy patients	169
4.2.2	Nasal gene therapy patients	170
4.3	Dose delivery and duration	171
4.4	Post dose follow up	172
4.4.1	Pilot A: 10ml lung dose/1ml nasal dose patients (n=3)	172
4.4.2	Pilot A: 20ml lung dose/ 2ml nasal non-invasive lung dose patients (n=3)	173
4.4.3	Pilot A: 20ml lung/ 2ml nasal invasive cohort (n=10)	174
4.4.4	Pilot B: dose cohorts	174
4.5	Summary post-dose results according to cohort	175
4.5.1	Clinical	175
4.5.2	Physiology	176
4.5.3	Inflammation	178
4.5.4	Post dose follow up to day 28	179
4.6	Post hoc comparison of lung dose cohorts	180
4.7	Adjuvant therapy: 'As required' and planned	181
4.7.1	Ibuprofen	181
4.7.2	Prednisolone	181
4.7.3	Paracetamol	182
4.7.4	Bronchodilators	182
4.8	Efficacy outcomes	183
4.8.1	CFTR molecular results	183
4.9	Serious adverse events	184
4.10	Discussion	185
4.10.1	Trial design	186
4.10.2	Safety outcomes	188
4.10.3	Efficacy	190
4.10.4	Adjuvant therapy	191
4.10.5	Conclusions	192
5	<i>Results: nasal PD measurement</i>	193
5.1	Pilot study nasal PDs	193
5.2	Pilot study nasal PD results	195
5.2.1	Nasal PD circuit integrity	195
5.2.2	Agreement between analysers	195
5.2.3	Nasal PD results for initial (1ml) nasal dose cohort	198
5.2.4	Final PD values for main (2ml) nasal dose cohort	200
5.2.5	Actual nasal dose (absolute volume) delivered and evidence of CFTR function	220
5.2.6	Comparison with other CFTR outcomes	220
5.2.7	Characteristics of patients with chloride response on post dose nasal PD	221
5.3	Novel approach to nasal PD	222
5.3.1	Measurements in the non-CF and CF nose	222
5.3.2	Further experimental work	224
5.4	Discussion	226
5.4.1	Feasibility of airway PD testing	227
5.4.2	Strengths of design	228
5.4.3	Limitations of nasal PD procedure and trial design	230
5.4.4	Absence of association between nasal PD and molecular results	231
5.4.5	Influence on chloride response	231
5.4.6	Influence on basal PD	232
5.4.7	Evidence of prolonged gene expression	232
5.4.8	Intermittent PD responses	233
5.4.9	Electrical changes and clinical correlates	234
5.4.10	Novel approach to nasal PD	235

5.4.11	Chapter summary and conclusions	237
6	<i>Results: Lower airway PD measurement</i>	239
6.1	Pilot study bronchoscopies	239
6.2	Lower airway PD	239
6.2.1	Agreement between JD and GD	240
6.2.2	Final lower airway PD results	241
6.3	Comparison with nasal PD results	247
6.4	Comparison with molecular <i>CFTR</i> results	248
6.5	Discussion	249
6.5.1	Experimental set up and design	249
6.5.2	Data analysis	251
6.5.3	Results from Pilot study	252
6.5.4	Chapter summary and conclusions	255
7	<i>Overall summary and conclusions</i>	257
7.1	Measuring outcome	257
7.1.1	Why?	257
7.1.2	When?	258
7.1.3	How?	258
7.2	The multidose trial	260
7.3	The future	261
8	<i>References</i>	262
9	<i>Appendix</i>	272
9.1	Run-In study: schedule of visits and investigations	272
9.2	Pilot study visits and investigations	273
9.3	Run-In study symptom score	274
9.4	Composition of nasal PD solutions	275
9.4.1	Ringers (R) solution	275
9.4.2	Ringers amiloride (RA) solution	275
9.4.3	Zero chloride amiloride (ZCA) solution	275
9.4.4	Zero chloride isoprenaline and Zero chloride amiloride isoprenaline solution	276
9.5	Composition of bronchial PD solutions	276
9.5.1	Ringers (R) solution	276
9.5.2	Zero chloride (ZC) solution	276
9.5.3	Zero chloride isoprenaline solution	276
9.6	Nasal PD analysis agreement for chloride secretion	277

List of Tables

Table 1.1. Summary of the safety outcomes from the previous placebo controlled gene therapy clinical trial in CF patients by Alton et al.	29
Table 1.2. Summary of the efficacy outcomes from the previous placebo controlled gene therapy clinical trial in CF patients by Alton et al.....	30
Table 1.3. Challenges of identifying appropriate patient population for CF gene therapy clinical trials.	34
Table 1.4. Outcome measures in the Pilot study categorised according to primary role in assessing safety or efficacy.	36
Table 2.1. Run-In study outcome measures described within this thesis	50
Table 2.2. Adjuvant medication administered with lung gene therapy dose.....	64
Table 2.3. Linear regression equations for calculation of FEV ₁ % predicted as used at start of Run-In and Pilot studies.....	68
Table 2.4. FEV ₁ % predicted values for the n=10 patients that crossed their 18th birthday during the first 4 visits of the Run-In study.....	73
Table 2.5. LCI data loss following SimpleWashout analysis and quality control review, according to Run-In study sites.....	82
Table 2.6. CT scoring grid used for analysis of Run-In chest CT scans.	87
Table 2.7. Adult and paediatric domains for the CF Quality of Life questionnaire (CFQ-R).....	89
Table 2.8. Questions in CFQ-R contributing to the Respiratory Symptom domain score.....	90
Table 3.1. Baseline characteristics of Run-In study participants.....	128
Table 3.2. Genotypes from Run-In study participants	129
Table 3.3. Comparison of visit 1 outcomes from patients with assay result at every one of the 4 study visits with the initial study population.....	131
Table 3.4. Summary results from chest CT scans. Summary results from n=178 patients at a single time point scored according to the methodology described in Chapter 2, section 2.4.3.....	137
Table 3.5. Mean inter-visit coefficient of variation (CV) between visits for the outcome of interest	140
Table 3.6. Correlations between airway physiology outcomes at study visit when CT carried out.....	143
Table 3.7. Correlations between component of CT score and remaining CT outcomes	145
Table 3.8. Correlations between physiology, inflammatory and quality of life outcomes with CT outcomes, at study visit when CT carried out.	148
Table 3.9. Correlations between physiology outcomes and measures of systemic inflammation and quality of life at study visit when CT carried out	150

Table 3.10. Correlations between quality of life and measures of systemic inflammation	150
Table 4.1. Baseline characteristics of the n=35 patients receiving a nebulised dose of CF gene therapy.....	170
Table 4.2. Characteristics of patients according to lung dosing cohort.	180
Table 4.3. Proportion of patients receiving additional medication on an ‘as required’ basis post dosing, according to lung dose cohort	181
Table 4.4. Vector specific CFTR mRNA results from nasal brushings for all patients in the nasal cohort.....	183
Table 4.5. mRNA results from bronchial brushings for all patients in the bronchoscopic cohort (Pilot part A).	184
Table 5.1. Schedule of nasal PDs for each patient receiving a 1ml or 2ml dose of CF nasal gene therapy	194
Table 5.2. Intra-reporter variability in nasal PD analysis.....	197
Table 5.3. Pre-dosing nasal PD summary for the 1ml nasal dose patients	198
Table 5.4. Pre-dosing nasal PD summary for the 2ml nasal dose patients.	201
Table 5.5. Proportion of patients with a chloride ‘response’ according to each definition and section(s) of nasal PD test at day 2 post dose	214
Table 5.6. Proportion of patients with a chloride ‘response’ according to each definition and section of nasal PD test at day 6 post dose	214
Table 5.7. Proportion of patients with a chloride ‘response’ according to each definition and section of nasal PD test at day 14 post dose.	215
Table 5.8. Proportion of patients with a chloride ‘response’ according to each definition and section of nasal PD test at day 28 post dose.	215
Table 5.9. Intrinsic chloride secretion ‘response’ post dose at any time point according to definition used	218
Table 5.10. Total chloride secretion ‘response’ post dose at any time point according to definition used	219
Table 5.11. Actual volume of nasal gene therapy delivered (2ml dose cohort) and nasal PD outcome for intrinsic and total chloride secretion	220
Table 5.12. Summary of nasal PD chloride secretion outcome and molecular evidence of vector specific CFTR on post 2ml dose nasal brushing.	221
Table 5.13. Patient characteristics according to evidence of chloride ‘response’ on nasal PD (2ml dose cohort).	222
Table 6.1. Intra class correlation coefficients (ICC)	241

List of Figures

Figure 1.1. Schematic diagram of ion transport in airway epithelial cells.....	21
Figure 1.2. Relationship between clinical studies in the UK CF Gene Therapy Consortium's clinical programme.	35
Figure 2.1. Lung dose nebuliser set up.	62
Figure 2.2. Difference in absolute FEV ₁ % predicted values for identical raw data according to reference source at Run-In visit 1.....	69
Figure 2.3. Bland Altman plot for absolute FEV ₁ % according to reference source.	71
Figure 2.4. Comparison of FEV ₁ % predicted and z scores.	72
Figure 2.5. Annual change in absolute % predicted FEV ₁ for the 10 patients transitioning across their 18th birthday during the Run-In study.	74
Figure 2.6. Illustration of T-piece flow past for LCI measurement.	76
Figure 2.7. Multiple breath washout trace illustrating wash-in and wash-out phases of the test.	77
Figure 2.8. Influence of Offset correction on LCI result.....	81
Figure 2.9. Nasal PD equipment set up.	91
Figure 2.10 Nasal PD traces in normal and CF subject.....	93
Figure 2.11. Photograph of adapted Foley catheter.	95
Figure 2.12. Testing electrode offset between two Ag/AgCl electrodes.....	112
Figure 2.13. Development of electrode interface for bronchial PD circuit: testing offsets with electrode embedded in agar.....	113
Figure 2.14. Development of electrode interface for bronchial PD circuit: testing offsets with electrode placed in layer of electrode cream on top of agar.	114
Figure 2.15. Schematic representation of initial bronchial PD set up.	116
Figure 2.16. Initial bronchial PD set up tested in the non-CF nose.	117
Figure 2.17. Schematic representation of the modified bronchial PD set up. ..	120
Figure 3.1. Number of patients (N) participating in each Run-In study visit. ...	127
Figure 3.2. Interval between study visits 1-4	130
Figure 3.3. FEV ₁ % predicted values and z scores across all 4 study visits.	132
Figure 3.4. Coefficient of variation in FEV ₁ % predicted according to visit 1 FEV ₁ z score.....	133
Figure 3.5. Summary data for spirometry outcomes at each study visit.	134
Figure 3.6. Summary data for lung clearance index (LCI) at each study visit. ..	135
Figure 3.7. Variability in LCI between study visits according to quartile of visit 1 LCI in Run-In population.	136

Figure 3.8. Blood markers of systemic inflammation.	138
Figure 3.9. CFQ-R scores for each questionnaire domain at Run-In visit 1.	139
Figure 3.10. CFQ-R Respiratory domain score summary data at each study visit.	139
Figure 3.11. Spirometry and LCI correlations.	142
Figure 3.12. Severity and extent of bronchiectasis on CT scan.	144
Figure 3.13. Extent of bronchiectasis and FEV ₁ , MEF ₂₅₋₇₅ and LCI.....	147
Figure 3.14. Change in FEV ₁ % predicted per year.	152
Figure 3.15. Annual change in FVC and MEF ₂₅₋₇₅ % predicted and ratio of FEV ₁ /FVC.....	153
Figure 3.16. Annual change in LCI	153
Figure 3.17. Annual change in blood total white cell count, neutrophils and C-reactive protein	154
Figure 3.18. Annual change in the Respiratory domain score of the CFQ-R	155
Figure 3.19. Annual change in FEV ₁ % predicted according to FEV ₁ z score at visit 1.....	156
Figure 3.20. Annual change in FEV ₁ and LCI.....	157
Figure 3.21. Annual change in Respiratory domain score of CFQ-R and FEV ₁ z score.....	158
Figure 3.22. Annual change in Respiratory domain score of CFQ-R calculated according to slope derived from 2 or more (rather than 4) study visits.....	159
Figure 4.1. Proportion of nebulised gene therapy complex administered according to dose cohort	171
Figure 4.2. Proportion of nasal gene therapy dose administered according to dose cohort.....	172
Figure 4.3. Temperature 8 hours post start of lung dosing according to dose cohort.	176
Figure 4.4. Maximum drop in FEV ₁ (as % of raw volume) on dosing day.....	176
Figure 4.5. Change in KCOc between baseline and day 2 according to dose cohort.	177
Figure 4.6. Change in LCI according to dose cohort.....	178
Figure 4.7. Change in CRP and white cell count post dose.....	179
Figure 4.8. FEV ₁ % predicted values and age at recruitment according to lung dosing cohort.....	180
Figure 5.1. Flow chart describing outcome of nasal PD data analysis and consensus review process between GD and JD.....	196
Figure 5.2. Intrinsic chloride secretion at all pre- and post dosing tests for the n=3 patients receiving a 1ml nasal dose.	199
Figure 5.3. Total chloride secretion at all pre- and post dosing tests for the n=3 patients receiving a 1ml nasal dose	200

Figure 5.4. Paired results for Basal PD and response to Ringers amiloride perfusion.	203
Figure 5.5. Intrinsic chloride secretion at all pre- and post dosing tests for the n=19 patients receiving a 2ml nasal dose.	207
Figure 5.6. Intrinsic chloride secretion at post dose for patients with tests performed at day 2,6,14 or 28	208
Figure 5.7. Total chloride secretion at all pre- and post dosing tests for the n=19 patients receiving a 2ml nasal dose.	212
Figure 5.8. Total chloride secretion at post dose for patients with tests performed at day 2,6,14 or 28	213
Figure 5.9. Chloride transport at post dosing time points for all patients with tests at day 2, 6, 14 and 28 post dose.....	216
Figure 5.10. Recovery to baseline PD following application of maximal clamp setting.....	223
Figure 5.11. Change in baseline PD in response to clamp setting	224
Figure 5.12. Silver/silver chloride ball electrode.	225
Figure 5.13. Comparison of electrode set up with LR5	226
 Figure 6.1. Data analysis algorithm for Pilot study lower airway PD results	240
Figure 6.2. Intra-patient variability in basal PD in the lower airway	242
Figure 6.3. Intra-patient variability in distal zero chloride isoprenaline (ZCI) response.	243
Figure 6.4. Site specific basal carinal measurements.....	244
Figure 6.5. Mean basal carinal and distal measurements.	245
Figure 6.6. Zero chloride responses for n=8 patients excluding the 2 outliers..	246
Figure 6.7. Zero chloride responses for all n=10 patients.....	246
Figure 6.8. Correlation between pre- and post dose bronchoscopy measurement of chloride secretion with mean carinal basal PD	247

Abbreviations

Abbreviations for clinical studies

Full study title	Abbreviated title use throughout thesis
Longitudinal assessment of clinical measurements in patients with cystic fibrosis in preparation for a clinical trial of CFTR gene therapy	Run-In study
Evaluation of safety and gene expression with a single dose of pGM169/GL67A administered to the nose and lung of individuals with cystic fibrosis	Pilot study
A randomised, double-blind, placebo-controlled Phase 2B clinical trial of repeated application of gene therapy in patients with cystic fibrosis	Multidose trial
Cationic lipid-mediated CFTR gene transfer to the lungs and nose of patients with cystic fibrosis: a double-blind placebo-controlled trial	Lancet study

Abbreviations used throughout the thesis

Ad	Adenoviral
AAV	Adeno-associated viruses
Ag/AgCl electrode	Silver/Silver chloride electrode
ATS	American Thoracic Society
cAMP	Cyclic adenosine monophosphate
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance regulator (gene)
CFTR	Cystic fibrosis transmembrane conductance regulator (protein)

Cl ⁻	Chloride ion
CMV	Cytomegalovirus
CpG	Cytosine-phosphate-Guanine
CRF	Case record form
CRP	C-reactive protein
CT	Computed tomography
CTEU	Clinical Trials and Evaluation Unit
DMPE-PEG5000	Polyethylene glycol containing lipid
DOPE	Diioleoyl-sn-glycerophosphatidylethanolamine
DSMB	Data safety monitoring board
ENaC	Epithelial sodium channel
ERS	European Respiratory Society
EVA	Ethyl vinyl acetate
FDA	US Food and Drug Administration
FEV ₁	Forced expiratory volume in the 1st second
FEV ₆	Forced expiratory volume in the first 6 seconds
FRC	Functional residual capacity
FVC	Forced vital capacity
GL67A	Cationic lipid used in clinical programme
GTAC	Gene Therapy Advisory Committee
Hb	Haemoglobin
ICC	Intraclass correlation coefficient
IL-6	Interleukin-6
IQR	Interquartile range
LCI	Lung clearance index
LLN	Lower limit of normal
LR4	Logan Research voltmeter model used for PD testing
LR5	Novel Logan Research voltmeter model
MBW	Multiple breath washout
MDI	Metered dose inhaler
MEF ₂₅₋₇₅	Forced expiratory flow between 25% and 75% of FVC
MHRA	Medicines and Healthcare products Regulatory Agency
MRSA	Methicillin resistant <i>Staphylococcus aureus</i>

Na ⁺	Sodium ion
PBS	Phosphate buffered saline
PPC	Percent positive cells
PD	Potential Difference
nPD	Nasal Potential Difference
R	Ringers solution
RA	Ringers amiloride solution
RhDNase	Recombinant human deoxyribonuclease
SAE	Serious adverse event
SD	Standard deviation
SeV	Sendai virus
SOP	Standard operating procedure
SUSAR	Suspected unexpected serious adverse reaction
UKCFGTC	UK Cystic Fibrosis Gene Therapy Consortium
ULN	Upper limit of normal
WCC	Total white cell count
ZCI	Zero chloride with isoprenaline solution
ZCA	Zero chloride amiloride solution
ZCAI	Zero chloride amiloride with isoprenaline solution

The following abbreviations have been used to identify the medical team (including the author) involved with the Run-In and Pilot study at the Royal Brompton Hospital site and are summarised in this thesis as follows:

GD	Dr Gwyneth Davies (author)
JD	Dr Jane Davies
EA	Professor Eric Alton
TB	Dr Tabinda Burney

Key members from the Edinburgh site referred to in this thesis included:

NB	Dr Nick Bell
HS	Dr Helen Sheridan
GM	Dr Gordon Murray

The following abbreviations have been used to identify the clinical sites described in this thesis:

RBH	Royal Brompton Hospital
RHSC Edinburgh	Royal Hospital for Sick Children, Edinburgh
WGH	Western General Hospital, Edinburgh

UK Cystic Fibrosis Gene Therapy Consortium Strategy Group

Members of the UKCFGTC Strategy Group are listed here according to Consortium site:

Imperial College London

Professor Eric Alton

Dr Jane Davies

Dr Uta Griesenbach

Ms Tracy Higgins

University of Edinburgh

Dr Chris Boyd

Dr Alistair Innes

Professor David Porteous

University of Oxford

Dr Deborah Gill

Dr Steve Hyde

1 Defining the context

1.1 Clinical evaluation of novel therapeutic agents

There is a constant search for better ways to treat patients with disease, irrespective of the underlying causality. This is particularly true for life-limiting diseases associated with significant morbidity. The processes involved in the journey from identification of a potential novel product to licensing and introduction in clinical practice result in only a fraction succeeding the translation from bench to bedside. Knowledge of the natural history of the disease of interest is imperative to ensure that an appropriate evaluation takes place. Even within a body system (e.g. respiratory), there is no single outcome measure to assess efficacy. Those optimal for a novel drug against community acquired pneumonia will be different to those for asthma. Consideration of severity of disease, and patient characteristics such as age, is also important (1;2).

This thesis describes an evaluation of outcome measures, both existing and novel, for a non-viral gene therapy product for cystic fibrosis (CF) developed by the UK Cystic Fibrosis gene therapy Consortium (UKCFGTC). The UKCFGTC was founded in 2001 as a collaborative approach in working towards gene therapy for CF with detectable clinical benefit. It includes clinical, scientific and administrative personnel from three geographical locations in the UK: Edinburgh, Oxford and London. Patients enrolled in the clinical programme attend either the Royal Brompton Hospital in London, or one of two centres in Edinburgh (Royal Hospital for Sick Children or Edinburgh Royal Infirmary).

The Consortium's clinical programme, which involves both observational and interventional studies, provides the platform for the data analysis presented.

1.2 Cystic fibrosis

Cystic fibrosis (CF) is a life-shortening, autosomal recessive, chronic multi-system disease, caused by mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene on chromosome 7. It is the commonest lethal inherited condition in Caucasian populations, affecting around 1:2500 births in the United

Kingdom and with a UK prevalence > 7000. Over 1500 mutations have been detected in the *CFTR* gene to date although in the UK p.Phe508del accounts for around 70% of cases (3). The clinical phenotype is difficult to predict according to underlying genetic mutation; patients with identical mutations can have widely contrasting disease courses (4;5).

1.2.1 *CFTR*

In health, the *CFTR* gene is responsible for the production of CFTR protein. CFTR functions as a cyclic adenosine monophosphate (cAMP) dependent chloride channel in epithelial cells as well as acting in a regulatory role for the epithelial sodium channel (ENaC) and the outwardly rectifying chloride channel (6). There is also evidence for a role in water conductance via aquaporin channels (7) and potassium ion transport at the basolateral surface of the airway epithelium (8). *CFTR* expression is thought to be maximally expressed in the submucosal glands and in the surface epithelium of the distal small airways (9;10).

In CF, defective *CFTR* function results in abnormal ion transport across epithelium with absent or impaired chloride (Cl^-) secretion and exaggerated sodium absorption (Figure 1.1).

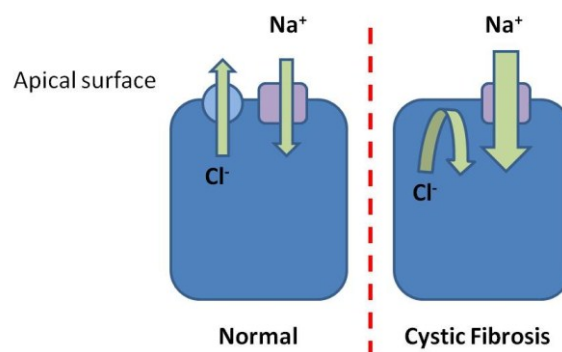


Figure 1.1. Schematic diagram of ion transport in airway epithelial cells. In healthy cells, CFTR (represented by pale blue circle) functions as a chloride ion (Cl^-) channel allowing chloride secretion and regulation of sodium ion (Na^+) transport. In cystic fibrosis, absent or defective CFTR does not permit chloride secretion and dysregulation of the sodium ion channel ENaC (purple square) results in sodium hyper-absorption.

The downstream consequences of the ion transport defects result in a reduced volume of airway surface liquid which impacts on the ability to achieve effective

mucociliary clearance (11;12). This disrupts the host's innate defence systems and provides opportunity for bacteria to thrive. Airway inflammation and chronic infection of the airway ensues. The downstream consequences of *CFTR* dysfunction ultimately lead to mucus plugging and airway remodelling, with airway destruction and extensive bronchiectasis can be detectable on chest computed tomography (CT) scan and obstructive ventilatory defects in tests of airway physiology.

1.2.2 Morbidity and current approach to management

Although respiratory morbidity is the greatest burden in CF, the disease impacts on multiple body systems. The digestive system is commonly involved, with pancreatic exocrine insufficiency affecting >90% of patients (3). Co-morbidities such as cystic fibrosis related diabetes are seen with increasing prevalence as patients survive to older ages. Pulmonary exacerbations are associated with a decline in respiratory function and preservation of respiratory status is important to minimise morbidity. Historically, management of CF has been supportive. To mirror the disease burden there is an associated heavy treatment burden. This includes agents to improve airway clearance (e.g. regular daily physiotherapy and nebulised mucolytics), antibiotics, nutritional support and often many other concurrent medications to counteract the consequences of this multisystem disease.

With appropriate supportive treatment, life expectancy has improved considerably over the past four decades such that infants born today are predicted to have a median survival > 50 years (13). Historically, management strategies have tackled the downstream consequences of the underlying molecular defect in *CFTR*. The potential for treatments that could instead cure or modify disease progression increased following identification of the *CFTR* gene over two decades ago (14;15). Such approaches include gene therapy and small molecule drugs, and are exciting in their theoretical potential to favourably alter progression of disease and reduce the considerable burden of treatment associated with CF care.

1.3 Gene therapy for cystic fibrosis

Gene therapy describes the approach whereby a gene is delivered to target cells and then transcribed *in vivo* into functional protein to modify the pathogenesis of disease or the consequences thereof. The disease of interest and associated

morbidity determine the organs and type of cells that are targeted by gene therapy. In CF, as the conducting airways are responsible for the greatest morbidity and mortality, the airway epithelial cells are the target (16). In the case of CF and the *CFTR* gene, insertion of a healthy copy into airway cells aims to result in subsequent transcription of CFTR protein (with insertion in the cell membrane) which is able to fulfil its multitude of functions including ion transport and air surface liquid homeostasis.

The target cells for the gene therapy agent along with the mode of delivery (systemic or local) in part determine the potential longevity of effect. For example, if a cell can divide and differentiate, an integrating transgene has potential for lifelong expression. This contrasts with strategies which either target relatively short-lived, terminally differentiated cells (e.g. airway epithelial cells) or do not involve integration of the delivered gene into the host cell genome. In the latter, the lack of integration means that even if the gene were to be delivered to dividing cells, it would not be expressed in successive generations.

Thus, the duration of gene expression will be determined both by the natural life cycle of the cells targeted and the characteristics of the gene therapy agent. As CF is a lifelong, chronic disease, sustained long-term gene expression will be required; this will only be achievable by correcting airway progenitor cells with an integrating vector or by using an agent able to be repeatedly administered.

As a single gene disorder, CF has been considered an attractive target for gene therapy. The *CFTR* gene was discovered in 1989 (14;15) and gene therapy seemed an obvious next step. Both *in vitro* and *in vivo* (in mice) proof-of-principle for CF gene therapy were quick to follow the identification of *CFTR* (17;18). The first human *in vivo* clinical trials followed shortly later (19-21). However, it was soon clear that effective CF gene therapy in man would be more challenging than anticipated for the reasons described below. Therefore despite initial widespread enthusiasm, with multiple researchers committed to taking the field forward, focus has reduced significantly over recent years. Around the globe there are now only a few research groups committed to developing and investigating the potential for CF gene therapy as a realistic future therapeutic strategy.

1.3.1 Vectors

For CF gene therapy to work, efficient delivery of normal *CFTR* to the airway cells is a crucial first step. To help with this process, the *CFTR* DNA is delivered into cells by different vehicles known as vectors. Those used commonly include genetically-modified viruses and non-viral agents such as liposomes. Whilst viruses are generally more efficient, the repeated delivery needed to achieve sustainable expression of *CFTR* has proved impossible to date due to the host immune response (22).

1.3.1.1 Viral vectors

From first principles, it is easy to see why viruses may be attractive vectors for gene therapy. Viruses have evolved to employ sophisticated methods of penetrating the body's defence and invading the target host cells. These features have been utilised to increase the efficiency with which the *CFTR* gene can be delivered to the airway. The success with which airway epithelium can be transfected with viral vectors has overall been less impressive than anticipated. However, *in vitro* studies have demonstrated that the proportion of cells that need to be transfected to normalise chloride Cl^- ion transport is as low as 6%; the mechanism underlying the normalisation of measured Cl^- ion transport with such a low percentage of cells transfected may relate to intercellular gap junctions (23). This contrasts with a requirement to transfect almost 100% cells *in vitro* to normalise sodium ion (Na^+) transport (24). Which one of these functions (and/ or others) requires correcting, and to what extent, is therefore an important consideration when assessing the outcomes of *CFTR* gene therapy trials. A trial of a small molecule potentiator drug (Ivacaftor, VX-770) which demonstrated evidence of correction of chloride secretion and improvements in lung function despite variable or no change in sodium transport lends some support to the notion that correction of chloride secretion may be sufficient for clinical benefit (25).

Viruses which have been investigated as candidate vectors for CF gene therapy include adenoviruses, adeno-associated viruses, Sendai virus and Lentiviruses (16). Adenoviruses and adeno-associated viruses have been assessed in early phase clinical trials. Sendai virus and lentiviral vectors remain in the pre-clinical setting at present. Adenoviral (Ad) vectors were the first viral vectors evaluated in man for

CF gene therapy (19;21). However, significant limitations were discovered with immune responses limiting efficacy after re-administration (26). Furthermore, the primary receptor for Ad, rather than being conveniently located on the apical cell surface, is on the basolateral surfaces and therefore relatively inaccessible from the airway lumen without an agent to break down the tight junctions (27).

Adeno-associated viruses (AAV) have also been extensively investigated including in clinical trials (16). They differ from Ad viruses in size, in addition to favourable low host pathogenicity. However, their receptors are serotype-specific and most are similarly scarce on the apical surface of the airway epithelium. As they are small viruses, large genes such as *CFTR* may be difficult to package; to overcome this, approaches have been designed including truncating or dividing the *CFTR* gene (28). Recent progress with viral vectors has included directed evolution of viruses (using selection based criteria) or molecular techniques to engineer desirable characteristics to improve their transfection efficiency. For example, an exciting recent approach has been to combine the most favourable characteristics of two different viruses, neither of which in isolation is ideally suited for CF. This is termed pseudotyping and is being investigated for Sendai virus (SeV) and lentiviruses (16). Sendai virus (SeV) is highly efficient at transfecting airway epithelial cells due to the presence of sialic acid and cholesterol receptors on their apical surface. Another advantage is its cytoplasmic expression, which removes the barrier of the nuclear membrane common to other approaches. High levels of gene expression have been reported but, in common with Ad and AAV viruses, it cannot be successfully re-administered. Lentiviruses are also highly efficient gene transfer agents. They have certain advantages over other gene therapy agents including the ability to transfect both dividing and non-dividing cells, and long-term stable expression; they also seem, almost uniquely amongst viruses to be repeatedly administrable (29;30). However, they naturally target haematopoietic cells and do not possess the surface proteins which recognise receptors on respiratory epithelia. Therefore neither lentiviruses nor SeV are ideal vectors in isolation, but combining characteristics of both may produce a more attractive viral vector. The UK CF Gene Therapy Consortium has recently generated a simian immunodeficiency virus (SIV) lentivirus pseudotyped with Sendai viral surface proteins. The SeV envelope packaging removes the usual limitations of the ability of SIV to bind to and enter airway epithelial cells. A single administration of this lentiviral vector has recently been shown to result in transgene expression for periods of up to 2 years in a

mouse and uniquely, in our experience, re-administration is possible and remains efficacious (30). Furthermore, in comparison to the non-viral gene therapy product that we are evaluating in the UKCFGTC clinical trial programme (described section 1.5), markers of gene expression are several hundred times greater. Although this does not automatically translate into sustained efficacious expression in man, such long lasting expression is encouraging and may ultimately result in a clinical product with greater clinical efficacy requiring infrequent administration. This is important to the CF patient population who have a high treatment burden with usual care alone.

Relatively few patients (none with CF) have received lentivirus-based gene therapy. As with any novel therapeutic agent, but particularly with any novel viral vector, safety is paramount. Caution is necessary, as lentiviruses integrate with the host cell genome. Integration into haematopoietic cells following gene therapy using retroviral vectors for severe inherited immunodeficiencies led to the development of leukaemias in some of the patients (31;32). Airway epithelial cells are very different from the rapidly dividing cells of the bone marrow, and lentiviruses are likely to be safer than retroviruses. However, the past experiences with integrating viral vectors will mean that progression to clinical trials should only be undertaken following extensive pre-clinical toxicology studies and consideration of appropriate CF patient groups for any ‘first-in-man’ testing.

1.3.1.2 Non-viral vectors

There are several non-viral approaches which aim to aid transfection and preserve functionality of the delivered gene to the airway cells, albeit with generally lower efficiency than achievable with viral counterparts. Non-viral approaches include cationic liposomes, compacted DNA nanoparticles and naked DNA (33). Cationic liposomes consist of cationic lipids that are usually mixed with cholesterol and dioleoylphosphatidyl-ethanol-amine. When combined with DNA they form particles sized 100-500 nm in diameter which can penetrate cell membranes and enter cells (34). The complex of DNA and liposomes is resistant to nuclease degradation, thus improving the success rates of gene delivery. The UKCFGTC is using the cationic lipid GL-67A (Genzyme Corporation) in its clinical trial programme. This was compared against other non-viral vectors and taken forward to clinical trial on the

basis of its preclinical efficiency. This lipid has also been used in a previous trial of CF gene therapy with a different CFTR plasmid (35).

Other non-viral approaches which have been investigated include DNA nanoparticles, where plasmid DNA is compacted with a peptide substituted with polyethylene glycol, and the cationic polymer polyethyleneimine (PEI) (33).

A major advantage of non-viral vectors would be the potential for repeat administration. In the nose there is some evidence that repeat dosing does not reduce molecular efficacy (36) although data are lacking with regard to the lower airway.

1.3.1.3 Plasmids for non-viral gene therapy

Just as the type of vector to deliver the gene to the cells of interest is important, the characteristics of the plasmid DNA used in non-viral gene therapy are also crucial to the potential for gene expression and ultimate success. In previous clinical trials of CF gene therapy using a liposomal vector with plasmid DNA, a post-dosing inflammatory response has been observed which is considered likely due, at least in part, to the presence of Cytosine-phosphate-Guanine (CpG) dinucleotides present in the plasmid molecule (37). Unmethylated CpGs occur commonly in bacterial DNA (occurring with a much lower frequency (and mostly methylated) in mammalian DNA) (38), and can be detected by the innate immune system as foreign to the host via Toll-like receptor 9 signalling (38). This activates signalling pathways resulting in the release of pro-inflammatory cytokines. Due to its amplification in *Escherichia coli*, CpGs on the plasmid *CFTR* DNA were considered likely responsible for the post dosing flu-like illness associated with rise in inflammatory markers seen when previously administered to man (35;39). Partial reduction of CpG content in plasmids only partially reduced the inflammatory effects following delivery by DNA-liposome complexes to the mouse lung (40). However, complete reduction of CpGs led to an absence of an inflammatory response in mice (41).

Modification of the plasmid DNA can also be undertaken to maximise its potential for gene expression. The plasmid used in the CF gene therapy clinical trial by Alton et al (35) described in detail in section 1.3.2 incorporated a viral (cytomegalovirus

(CMV)) promoter, which had a rapid onset of action but was limited by a short lived duration of expression. In contrast, the *CFTR* plasmid DNA being investigated by the UKCFGTC in its clinical trial programme (pGM169) has been identified on the basis of extensive pre-clinical testing. It includes plasmid DNA which has been stripped of all CpG dinucleotides, and includes a promoter which has been optimised (Human Elongation Factor 1 α) to enhance duration of expression. In mice, with this product, sustained expression for over a month can be achieved with a single dose (41).

1.3.2 Clinical Trials

Since the early 1990s, there have been over 20 previous early phase clinical trials of CF gene therapy, which have involved dosing to the nose, lung or both (42). These trials have largely been designed to assess safety and molecular (e.g. *CFTR* mRNA) or functional (electrophysiological tests e.g. airway potential difference measurements) proof-of-principle for gene transfer and expression in the host. Variable outcomes for both safety and efficacy have been reported but the principle of *CFTR* gene transfer has been proved, both with viral and non-viral approaches (16).

A placebo controlled trial by Alton et al (35) of a single dose of non-viral nebulised CF gene therapy is referred to frequently throughout this thesis. Published over a decade ago, it represented a benchmark study for the field of non-viral CF gene therapy and provided the stimulus for continued development of a clinically effective CF gene therapy product by the UKCFGTC. For ease of reference it is referred to as the Lancet study throughout this thesis and outcomes described here are for comparison with the single dose CF gene therapy clinical trial in the UKCFGTC clinical programme, (the 'Pilot' study; see section 1.5.2). The Lancet study was a double blind, placebo controlled study involving 16 adult male patients with CF, with FEV₁ at recruitment >70% predicted. It aimed to investigate safety and functional evidence of efficacy. Eight patients received a single dose of the nebulised and nasal CF gene therapy product (*CFTR* plasmid including a CMV promoter, complexed with GL67/DOPE/DMPE-PEG5000 (the lipid component collectively referred to as GL67A)). The 8 patients randomised to placebo received the lipid component alone.

Outcome measures for the Lancet study included measures to assess safety and efficacy. Patients were seen for a single pre-dose visit which included bronchoscopy and then for serial visits post-dosing at time points determined by the study protocol. Lung and nasal doses were administered at separate visits with an interval of 6 days. Safety outcomes included clinical assessment, chest computed tomography, lung function, sputum analysis (including sputum induction) and histological analysis. Efficacy was assessed in the upper and lower airway with molecular tests (CFTR mRNA and DNA), functional measurements of CFTR (upper and lower airway PD measurements), and in chloride efflux and bacterial adherence assays. All post-dosing bronchoscopies were performed at day 2. Patients were followed up until day 29 post lung dose. Post administration of the lung dose, patients in the active group experienced a self-limiting influenza type illness which resolved within 2 days. The safety outcomes and key results are summarised in Table 1.1.

Outcome measure	Mean maximal change from baseline in active treatment cohort	Detection of maximal change (time point)
FEV ₁	-17% (placebo: -15%)	Within 6 hours post dose
FVC	-17% (placebo: -15%)	Within 6 hours post dose
Gas transfer	-7.5% in active group	Day 2. Normalised by time of next measurement at day 29
C-reactive protein	Rise from 8.5 to 42.5mg/L	48 hours. Returned to baseline by day 6
Serum IL-6	Rise from 6 to 25 pg/mL	6 hours post dose
Total white cell count	Rise from 9.0 to 12.1 x10 ⁹ /ml	6 hours post dose
Chest computed tomography (CT) scan	No significant change at day 2 or 29	N/A

Table 1.1. Summary of the safety outcomes from the previous placebo controlled gene therapy clinical trial in CF patients by Alton et al. The table summarises the results from airway physiology, systemic inflammatory and structural outcomes from the previous clinical trial of non-viral CF gene therapy by Alton et al (35) including magnitude and time point of maximal change. Time points are post lung dose administration; FEV₁ forced expiratory volume in 1 second, FVC forced vital capacity.

The *CFTR* molecular and functional efficacy data from the Lancet study is summarised in the table below (Table 1.2), which has again been summarised to allow comparison with the results from the Pilot study.

Outcome measure	Mean change (active) group	Post dose tests
Bronchial brushings	No vector specific <i>CFTR</i> mRNA.	Day 2 (single post dose measurement)
Nasal brushings and nasal biopsy	No vector specific <i>CFTR</i> mRNA	Day 2 and 4 post dose
Lower airway potential difference (PD)	No significant change in basal PD or response to amiloride in either active or placebo group. Increase towards normal values of chloride secretion in 6/8 patients in active group. Significant increase in active but not placebo group ($p < 0.05$).	Day 2 (single time point post dose)
Nasal PD	No significant change in basal PD or response to amiloride in active or placebo group. Significant difference between active and placebo groups in intrinsic chloride secretion ($p < 0.05$). No change in response to isoprenaline in either group	Follow up (6 tests) post nasal dose to day 23

Table 1.2. Summary of the efficacy outcomes from the previous placebo controlled gene therapy clinical trial in CF patients by Alton et al. As per safety outcomes, the table summarises outcomes including post dose time points for data collection from the previous clinical trial of non-viral CF gene therapy by Alton et al (35). Time points are post lung administration unless otherwise stated.

In addition to confirming proof-of-principle for *CFTR* gene transfer with molecular or electrophysiological tests, general themes which have emerged from the Lancet study and other clinical trials to date include a reduction in gene expression after subsequent doses of viral gene transfer agents, and that both viral and non-viral

approaches may induce an inflammatory response. Few studies have investigated clinical end points, and when included have largely been those which were not powered sufficiently for such end points (having been designed to assess primarily safety). A reduction in sputum inflammatory markers has been reported in studies involving viral (AAV) and non-viral (cationic liposome) vectors (35;43). Small but significant improvements in lung function were reported in a CF gene therapy trial involving an AAV vector but this was not confirmed in a subsequent, larger trial (43;44).

1.3.3 Challenges

The airway has proved a challenge for effective topical delivery of gene therapy. The ciliated epithelial cells lining the airway are generally considered the target for topical application via nebulisation yet these cells, which are key in the lung's defence mechanisms, are specifically designed to keep foreign particles out. The requirement for *CFTR* expression in other cell types in the airway is at present unknown. This may be a particularly relevant issue for cells which may be harder to reach with nebulisation including those within submucosal glands (9), and also the putative airway stem cells (such as basal cells). There are also physical challenges with the accessibility of the distal airways which are considered fundamental to disease pathology in CF, and the presence of thick sticky secretions lining the airway walls. The technology of different commercially available nebulisers to deliver the gene therapy has been an important consideration in accessing these distal airways, including the stability of the product during nebulisation and resultant particle size. For example, aerosolisation of DNA alone has been found to result in loss of function *in vitro* (45).

Sustained efficacy has been a key challenge to CF gene therapy. Gene therapy is attempting to overcome well-evolved lung defence mechanisms. A key obstacle is to negotiate the body's own defences: plasmid DNA or engineered viruses may be seen as foreign by the host. Minimising any host immune response is therefore important to allow repeated delivery of the gene therapy. A further challenge is for the DNA to pass through the cell membrane, evade the cytoplasmic defence mechanisms and penetrate the nuclear membrane. Finally, the challenge of measurement and outcome should also not be overlooked. Quantifying the distribution and functional efficacy of CF gene therapy *in vivo* is challenging for

early clinical trials which do not aim to look for clinical benefit. If a CF gene therapy product is to ultimately become a viable clinical option, it must fulfil a number of requirements. The current challenge for CF gene therapy clinical trials is to demonstrate clinical efficacy. Important considerations for CF lung gene therapy include:

- Clinical efficacy demonstrated with appropriate outcome measures
- Sustained clinical efficacy with repeat administration
- Acceptable side effect profile

1.4 Why we need ‘better’ outcome measures for CF clinical trials

Novel medicinal products are required to pass through preliminary clinical studies in man before progressing to later phase trials. Initial clinical studies require evaluation of safety (with or without an assessment of efficacy) in a limited number of patients. At this early stage of clinical evaluation, efficacy may be evaluated at a molecular or cellular level indicative of functional change (“does it work?”) rather than clinically detectable endpoints (“does it help the patient?”) such as forced expiratory volume in one second (FEV₁). A multidose study involving considerably more subjects must show clinical efficacy. Therefore, the optimal outcome measure to evaluate *in vivo* ‘success’ evolves during the evaluation for any novel therapeutic agent (2).

Early death is of most concern to CF patients, but the improved prognosis of CF means that this is not a practical clinical trial outcome. The primary end-point of CF clinical trials has most commonly been spirometry, and in particular FEV₁. However, the slow annual decline (around 1% per year) in FEV₁ means that adequately powered studies need large numbers of patients to detect a change in the rate of decline (46). Furthermore FEV₁ lacks sensitivity to detect early structural CF lung disease (47-49). This means we need new surrogate end-points for clinical trials. These will vary with the severity of the disease and the stage of drug evaluation. Importantly, the sensitivity, prognostic ability and variability of

any novel surrogate must be known. One outcome measure gaining prominence is lung clearance index. This has recently been shown to be superior to FEV₁ in detecting a treatment effect of a small molecule drug in patients with early stage CF lung disease (50). There is also interest in the use of patient reported outcomes as methods of detecting clinical efficacy, but there is a lack of consensus amongst the regulatory authorities regarding their role. It is also important that any outcome measure used in clinical trial is recognised by the relevant regulatory authority. The US Food and Drug Administration (FDA) and European Medicines Agency require new therapies to be assessed using FEV₁ as a primary endpoint. However, there is increasing recognition that this may not be suitable for all studies.

There is a balance between opportunity for effective delivery of CF gene therapy and potential to detect a change in outcome. For example, a potential disease modifying treatment for young children may be easy to deliver by nebulised therapy but difficult to assess efficacy over the short to medium term due to sensitivity of existing assays in early disease. Conversely adults with late stage disease may have airways plugged with thick mucus which hinders delivery, yet abnormal baseline outcome measures (e.g. in spirometry) in which a treatment effect could easily be detected. These considerations are summarised in Table 1.3. Ensuring that we use the best available tests and have confidence in interpreting the results is not only important in the context of clinical trials but also for improving standards in routine clinical care.

Infants and young children		Older children and adults	
Potential advantages	• Ease of delivery and access to airway surface • Ability to 'prevent' or reduce airway disease	Potential advantages	• Opportunity to measure short to medium term efficacy due to baseline abnormal respiratory status • Subjects can consent to participation in clinical trials
Potential disadvantages	• Difficult to measure short to medium term efficacy • Ethics of novel treatment with unknown risks when expected clinical course in individual unknown	Potential disadvantages	• Physical barriers to effective delivery e.g. airway secretions • Apparent negative results due to physical barriers and irreversible pathology may lead to rejection of novel therapy

Table 1.3. Challenges of identifying appropriate patient population for CF gene therapy clinical trials.

1.5 The UK CF Gene Therapy Consortium clinical programme

The UKCFGTC aims to develop gene therapy for CF with clinically detectable benefit. It is evaluating a non-viral CF gene therapy product in its current clinical programme. Three clinical studies comprise what is commonly referred to as 'Wave 1' of this programme; the Run-In study, the single dose 'Pilot' study and a double blind placebo controlled multi-dose clinical trial. The relationship between the Run-In, Pilot and multidose trial in the Consortium's Wave 1 programme and the aim(s) of each study are summarised in Figure 1.2.

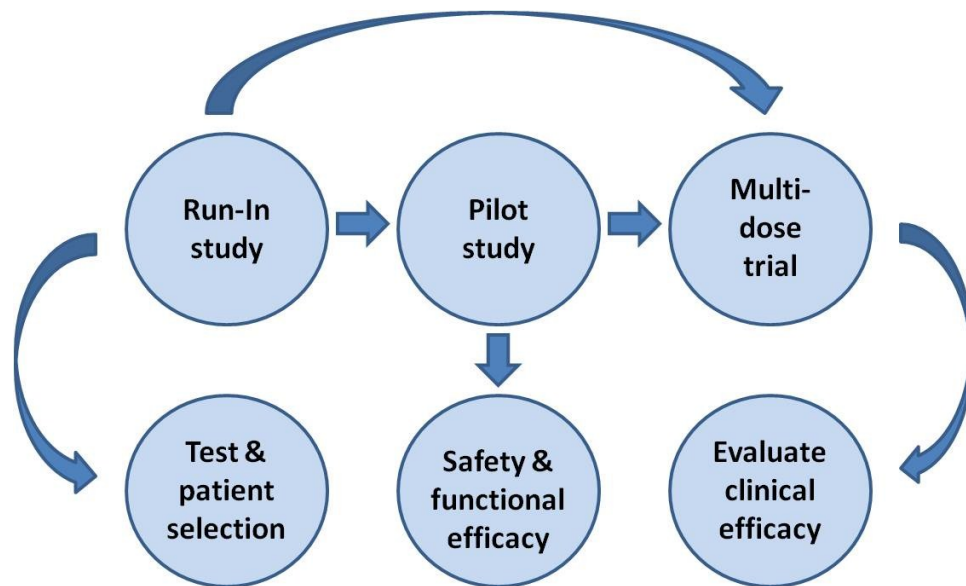


Figure 1.2. Relationship between clinical studies in the UK CF Gene Therapy Consortium's clinical programme.

A description of the Run-In study and Pilot study are given below. Recruitment to the multidose trial was in progress at the time of writing this PhD and therefore results from this study are not included within this thesis.

1.5.1 The Run-In study

The Run-In study was a longitudinal observational study involving patients with CF recruited in London and Edinburgh as part of the UKCFGTC clinical programme. The full title of the study is 'Longitudinal assessment of clinical measurements in patients with cystic fibrosis in preparation for clinical trial of CFTR gene therapy', referred to as the 'Run-In' study throughout this thesis.

By collecting serial measurements from a wide range of tests during periods of clinical stability, it aimed to identify both the most appropriate outcome measures and the most appropriate patients to take forward to multidose trial. Although the Run-In study was purely observational, outcome measures reflected interest in the potential to effectively deliver and detect clinical efficacy of a nebulised gene therapy product in a future interventional trial. They included outcomes encompassing particle deposition, airway physiology, airway structure, systemic and airway inflammation, and patient reported outcome measures. A list of outcome measures investigated within the study can be found in Appendix 9.1.

1.5.2 The single dose ‘Pilot’ study

A phase 1b clinical trial of non-viral CF gene therapy has been undertaken by the UKCFGTC within the time course of this PhD. The study is entitled ‘Evaluation of safety and gene expression with a single dose of pGM169/GL67A administered to the nose and lung of individuals with cystic fibrosis’. It enrolled patients with CF aged 16 years and over with a forced expiratory volume in one second (FEV₁) > 60% predicted at recruitment. Patients were given a single nebulised and/ or nasal dose of CF gene therapy (pGM169/GL76A). For simplicity and consistency the study is referred to as the ‘Pilot’ study throughout this thesis.

The primary aim of the Pilot study was to investigate safety and choose a dose to take forward into the multidose trial. It was also designed to take the opportunity to explore gene expression and *CFTR* function and the end points used in the trial reflected both of these aims.

Outcome measures for the Pilot study reflected the trial design of assessing safety and evidence of molecular and functional efficacy (Table 1.4). The Pilot study was not designed to investigate clinical efficacy.

Safety	Efficacy
Clinical examination and pulse oximetry	<i>CFTR</i> mRNA
Whole blood and serum markers	Upper airway PD
Spirometry	Lower airway PD
Gas transfer	Sputum inflammatory markers
Lung clearance index (LCI)	
Chest CT scan	
Bronchial biopsy	

Table 1.4. Outcome measures in the Pilot study categorised according to primary role in assessing safety or efficacy.

A full list of trial outcome measures is given in Appendix 9.2. This includes many established outcome measures in CF clinical trials (e.g. spirometry), but also others with as yet uncertain clinical roles.

1.6 Outcome measures of interest

The work of this PhD thesis is an evaluation of the outcome measures used in both longitudinal study (Run-In study) and an interventional clinical trial for CF (Pilot study), as well as an investigation into the development of novel outcome measures relevant to the field.

There are two main themes in this thesis. The first involves physiological measurements in the lower airway (e.g. spirometry) and their relationship with other outcomes (e.g. structural outcomes as measured by computed tomography). The other involves a functional readout at the cellular level of ion transport across epithelial cells (airway Potential Difference (PD)). These themes apply to opposite ends of the clinical trial process, but illustrate issues relevant to any phase of development.

1.6.1 Airway physiology outcomes

Physiological measurements of respiratory status including spirometry and lung clearance index represented important outcome measures for both the Run-In and Pilot studies. Longitudinal measurement of airway function in the Run-In study has offered a unique opportunity to investigate variability over time during apparent clinical stability. Serial measurements within patients pre and post nebulised gene therapy in an interventional study allow investigation of how post-intervention measurements relate to usual variability over time, and in particular how change in one test parameter reflects change in another physiological test of respiratory function.

In the Run-In study, airway physiology tests included spirometry and lung clearance index. In the Pilot study, they additionally included transfer factor of carbon monoxide (referred to as ‘gas transfer’ throughout this thesis).

1.6.1.1 Spirometry

Spirometry involves measurement of the flow and volume of inspired or expired air over time (51). In a forced expiratory manoeuvre, patients inhale to their maximum capacity (total lung capacity) from an end tidal expiration before

performing a forced manoeuvre in which they exhale rapidly with sustained expiration until no further air can be expired (residual volume). Spirometry indices, in particular forced expiratory volume in one second (FEV_1), are widely used in the clinical and research arenas for monitoring disease progression. They play an important role as both safety and efficacy measurements and are well established outcome measures in clinical trials (1;2). FEV_1 is attractive because the respiratory manoeuvre is relatively quick to perform, the test is non-invasive, test equipment relatively inexpensive and data collection in large epidemiological or cross-sectional studies enables comparison of disease populations against 'normal' values. However, as described previously it lacks sensitivity to detect early structural change in mild CF lung disease (47-49), and in clinical trials for CF the relatively small annual decline in % predicted values per year (around 1% per year) (46) necessitates large numbers of participants to achieve adequate power to detect any influence on this. Despite these limitations, it is currently the only CF efficacy outcome acceptable to regulatory authorities and as such, is the primary outcome in the vast majority of phase 3 clinical trials.

In comparison to FEV_1 , the forced expiratory flow between 25-75% of FVC (MEF_{25-75}) has been considered a measure of peripheral airways obstruction (52). Histological examination of tissue from patients with CF including organs from children undergoing lung transplantation has confirmed that there is extensive involvement of the small airways (53), and small airways disease is apparent on CT scans as areas of air trapping on expiratory images (54). Having outcomes which are able to detect changes according to geographical site (e.g. central large vs. small peripheral airways) is attractive. Although challenged by a known greater variability in measurement than FEV_1 , measurement of forced expiratory flow at 75% of FVC has been used as a primary outcome measure in clinical trial. For example, in a study of dornase alpha (recombinant human deoxyribonuclease (RhDNase)) delivered with nebulisers targeting either the small or larger airways, significant differences between baseline vs. 4 weeks were detected with this measure and also MEF_{25-75} but not with FEV_1 (55). MEF_{25-75} correlates with the ratio of FEV_1/FVC however in mild disease (i.e. mild obstruction to air flow) the former is associated with a greater decline (56). In practice the geographical region of the airway represented by the different spirometry outcomes is not uniformly agreed, and there is some evidence in the adult literature that MEF_{25-75} correlates poorly with other measures thought to reflect air trapping including FVC and the

plethysmographic derived value of residual volume divided by total lung capacity (57).

In studies involving only adult subjects with spirometry as an outcome measure, absolute change in expired volumes or flows can be compared pre and post intervention, but this is more complicated when children and young adults are involved, as lung volumes increase with height. To take account of the influence of height, age and gender in interpreting spirometry, there are a large and constantly expanding number of linear regression reference equations reported in the literature for comparing the raw expired volumes and flow rates to those expected (58-62). It is important that the reference source used is valid for the study population of interest. In the UK, it is currently recommended by the British Thoracic Society that Rosenthal values (62) are used for converting absolute values into % predicted values in childhood. In adults the European Community for Steel and Coal (ECCS) equations by Quanjer et al (61) are commonly used. Both use linear regression equations to predict values with which measured values can be compared.

However, these regression equations have limitations. For example, they do not recognise the non-linear relationship between lung volumes and height in adolescence (63). The original paper describing the Rosenthal reference equations does include factors for pubertal correction which may go some way to address this, but in usual clinical practice these are not taken into account. Stanojevic et al (64) have created a reference source that combines data from multiple studies spanning all ages in an effort to create greater accuracy and relate the relationship between observed and expected values with standard deviation score (z scores) in addition to % predicted values. This uses a method described by Cole (65) and now well established for growth charts where z scores are calculated according to values in 'look up' tables using a LMS approach (L=skewness; M= median; S= generalised coefficient of variation) which take account of distribution in the measurement and normalizes it. This is achieved by expressing the changing distribution with three curves representing the median, coefficient of variation and skewness allowing the calculation of z scores (66). The data on which the Stanojevic reference source is based are from four large studies of spirometric indices including the UK Rosenthal data and NHANES III data from the US (58;62). The implications of external reference sources have been investigated by the

author and are described within this thesis, including the differences in outcome according to whether discontinuous (separate adult and child) or continuous ('all ages') sources are used.

1.6.1.2 Lung clearance index

Lung clearance index (LCI) is a measure of ventilation inhomogeneity obtained using a multiple breath washout (MBW) technique. This is achieved either by 'washing out' nitrogen resident in the airways with 100% oxygen, or by measuring the progressive elimination of a non-absorbable, inhaled test gas over serial breaths from a starting equilibrium in the inspired and expired test gas concentration. From this equilibrium, the volume of expired gas taken to result in a concentration of the test gas $1/40^{\text{th}}$ of the starting concentration is measured (an arbitrary cut off based on the limits of nitrogen detection in the original mass spectrometry equipment used to develop the MBW technique). The LCI can be calculated from this cumulative expired volume. Early studies with LCI used measurements collected with a mass spectrometer, which is still considered the gold standard for testing (47;49;67). There has been increasing interest in LCI as an outcome measure in CF over the past decade. Abnormalities in LCI in young children with CF in comparison to controls in the presence of often normal spirometry demonstrated its use as a more sensitive indicator of early lung disease (47;49).

In some centres, all measurements continue to be performed using mass spectrometry although, increasingly, other methodologies are employed. The UKCFGTC uses the technique described by Horsley et al (68), with a modified Innocor device (Innovision, Odense, Denmark). LCI is being increasingly used as an outcome measure in longitudinal studies in early childhood and is becoming established as an outcome measure in studies involving CF patients with mild disease (69;70). Use as an outcome measure in CF clinical trials for older children and adults with pre-existing abnormalities in spirometric parameters including variability over time is at present unknown.

1.6.2 Airway potential difference (PD)

Abnormal or absent CFTR protein expression in the airway is associated with an ion transport defect that can be detected *in vivo*. This creates an opportunity to measure *in vivo* the function (rather than the presence) of CFTR protein, albeit indirectly. *De novo* CFTR protein (as could occur following gene therapy) or endogenous protein which has been potentiated or corrected can be detected by changes in the electrical characteristics of the tissue under test. Therefore it is not surprising that airway potential difference (PD), and particularly nasal PD, has an established role as an outcome measure in early phase clinical trials of novel agents for CF directed at tackling the underlying molecular defect. Nasal PD tests are only performed at a limited number of UK centres, a reflection on the equipment and technical expertise required for testing and a relatively recent evolving role in 'difficult' diagnosis. However the attractiveness of nasal PD as an outcome measure for clinical trials and as an aid in patients in whom the diagnosis of CF is difficult to confirm has led to its adoption by multiple centres on a worldwide basis.

1.6.2.1 CFTR and its role in ion transport

As described in section 1.2.1, CFTR is a cyclic adenosine monophosphate (cAMP)-dependent chloride channel which also regulates sodium ion transport via the epithelial sodium channel (ENaC)(71;72) and in CF the ion transport defects involving chloride and sodium flux result in an altered potential difference (PD) across the respiratory epithelium, measurable in millivolts. When compared to healthy cells expressing CFTR, in CF the absent chloride secretion and sodium hyperabsorption result in an apical cell surface which is electrically more negative under basal or resting (non-stimulated) conditions (73). Amiloride, as a sodium channel blocker, results in a greater reduction (depolarisation) in nasal PD for CF than non-CF subjects (74;75) and subsequent ionic and pharmacological challenges to stimulate chloride secretion (via chemical gradients) are ineffective in CF but lead to pronounced hyperpolarisation in the presence of functional CFTR. There is recent evidence from a porcine model that abnormal transport of chloride ions (rather than abnormal transport of sodium ions) may be able to explain all of the abnormal airway PD measurements which are found in CF (76). Evidence in support

of this has since also been demonstrated in an *in vitro* model of cultured lower airway epithelial cells (77).

In vivo measurements of PD across the respiratory epithelium are made either in the upper (nose) or lower (bronchial) airway. Nasal PD detects changes in electrical potential at the airway surface, measured using a millivolt meter in real time whilst a series of solutions are perfused onto the airway surface. The perfused solutions allow an evaluation of the cells within any fluid contact for their response(s) dependent (as least in part) on CFTR function. Measurement of nasal PD *in vivo* was first described in healthy adult subjects by Knowles et al in the early 1980s (74;78). By measuring PD at the inferior surface of the inferior turbinate, he found no overlap between CF and non-CF subjects in basal PD measurements (74). Modification of the technique was performed by Alton et al, with measurement instead along the floor of the nasal cavity using an adapted Foley catheter containing diluted electrode gel and a solid silver/silver chloride (Ag/AgCl) electrode (79;80). The same group (Middleton et al) then described a protocol for measuring nasal PD with a series of perfusion steps including those permitting measurement of baseline, response to amiloride, and chloride secretion (75). The principles of this method were those employed in the UK CF Gene Therapy Consortium Pilot study.

Several groups have attempted to define normal ranges for nasal PD, although differences in experimental set up contribute to variable results. In an early report of measurements undertaken on the nasal floor, Alton et al measured basal PD in a CF and control population (80). In healthy controls, mean PD was -19 (95% CI -7.6 to -30.4). In CF, mean basal PD was -46.1mV (95% CI -27.7 to -64.5). Airway PD *in vivo* has been used to discriminate between CF and non-CF subjects (74;75), and correlates with the number and severity of *CFTR* mutations (81). In CF, the response to solution perfusion assessing response to sodium blocking agents or to encourage chloride secretion relates to clinical status (82). Nasal PD is included in the diagnostic algorithm for CF by the European Diagnostic Working Group for cases where diagnostic results are borderline or inconclusive (i.e. sweat chloride and *CFTR* genotyping) (83).

Additionally, airway PD measurements have been used as an outcome measure in clinical trials (35;36;84;85). Usually this involves nasal PD, but lower airway PD measurements have also been used in a small minority (35).

1.6.2.2 Airway PD as an outcome measure in CF gene therapy clinical trials

As a functional read-out of CFTR, nasal potential PD measurement has become an established end point for early phase clinical trials involving gene therapy.

Although the lung is the primary target, the nose has gained an important role in assessing molecular and functional evidence of gene transfer in early stage clinical trials of therapies targeting the underlying molecular defect largely with nasal PD. Measurements pre- and post gene therapy can detect any normalisation of values as a result of *de novo* CFTR expression and can provide information on the duration of expression. Due to the technical challenges and the lack of repeatability, lower airway PD measurements have been used less often. However, a restoration of 25% towards normality of electrical values in the basic underlying (chloride) ion transport defect in the lung was demonstrated in the Lancet study (35). The correlation between degree of electrical normalisation and clinical outcome is unknown.

In trials involving nasal dosing, nasal PD has been used as a main outcome for efficacy (20;35;36;42;86-91). However, nasal PD results have been variable, and correlation of any change in nasal PD with molecular evidence of *CFTR* has often been lacking (20;35;89). Looking at the PD results on a per patient basis (i.e. in isolation of group results), several suggest change towards non-CF values (20;35;36;87;89-93) however the responses have not been universal and the reasons for this are unclear. Many studies have suffered from limited numbers of pre-intervention tests. Within-subject variability in nasal PD has been reported within the literature and is high (94). Therefore it can be difficult to confidently interpret post treatment values if only a single pre-treatment nasal PD has been conducted. This has been the case in much of the published literature. Where changes in PD have been reported following gene therapy, these have generally been restricted to sections of the test which represent improved chloride transport rather than correcting sodium hyperabsorption. This is likely to reflect the higher proportion of cells required to be corrected before changes in sodium hyperabsorption can be

corrected (>90% vs. <10% for restoration of chloride transport) (23;24;95) and is a reason why improved methods of quantification are required. A PD measurement set up which had the capacity to quantify the success of *CFTR* correction would allow comparison with clinical efficacy or evaluation of response to therapy over time, and be advantageous in comparison to conventional methodology.

Use of nasal PD as an outcome measure is not limited to gene therapy evaluation, and recent studies have included trials of small molecule drugs (96-98). Not only does inclusion of nasal PD in such studies provide evidence of proof of principle, it could also allow retrospective analysis of how changes detected at the level of ion transport relate to the clinical picture and downstream consequences of disease.

It is not yet known how any change between pre- and post treatment results (in mV) relates to clinical outcome. In a phase II clinical trial (97), the small molecule potentiator drug Ivacaftor, led to a modest 3.5mV change in chloride secretion on nasal PD (chloride response); in contrast, sweat chloride in response to this orally-administered agent, fell dramatically by a mean of > 50 mmol/l and in some subjects into the normal range. The subsequent phase III clinical trial (in which nasal PD was not performed) reported a large change in absolute % predicted FEV₁ from baseline of 10.6% in comparison to the placebo group and a significant reduction in pulmonary exacerbation rate (99). Clearly therefore, small apparent changes in nasal PD can co-exist with much larger changes in clinically relevant, pulmonary end-points.

Present methods for PD measurement cannot determine the *quantitative* success of gene therapy *in vivo* because correction of a higher proportion of cells within epithelial sheets does not result in further ‘super correction’ of Cl⁻ ion transport (23), related to the arrangement of epithelial cells as if in a parallel electrical circuit. The clinical relevance of correcting a certain proportion of cells is unknown. The ability to quantify correction would allow comparison with clinical efficacy and could be used as a tool to assess the success of delivery over time (e.g. following repeated doses within the same patient to determine whether transfection occurs in the same or different cells). This would be potentially useful also for other novel CF treatments aimed at tackling the underlying molecular defect (and for which early phase clinical trials already use ‘standard’ nasal PD measurements as an outcome measure of functional efficacy).

1.7 Hypotheses, aims and objectives

1.7.1 Hypotheses

1. Non-viral *CFTR* gene transfer will lead to:
 - a. Expression of functional CFTR protein
 - b. Detectable improvements in airway health
2. Currently used methods to detect these changes may be optimised to:
 - a. Enhance sensitivity and allow quantification
 - b. Facilitate an understanding of geographical effects
3. Understanding the relationship between CFTR electrophysiological effects and the downstream consequences on airway physiology will allow the more rational design of future trials.

1.7.2 Aims

1. To optimise proximal and distal electrophysiological measures of *CFTR* function.
 - a. To obtain proximal (nasal) and distal (lower airway) potential difference measurements using conventional techniques modified to be compliant with current regulatory requirements
 - b. To design a novel electrophysiological method to allow quantification of corrected cells in the upper airway.
2. To refine both the methodology and analysis of measures of airway physiology, both conventional (spirometry) and more novel (lung clearance index).

3. To explore correlations between electrophysiological and physiological outcomes in a trial of CFTR gene therapy.

1.7.3 Objectives

The objectives are divided according to the two main themes of this thesis; namely airway physiology and airway potential difference (PD) measurement.

1. Airway physiology

- a. Analyse serial spirometric measurements focussing on the influences of variability and the impact of reference equations
- b. Compare the more novel measure, Lung Clearance Index, with conventional spirometry to assess its potential superior sensitivity and whether the two techniques are providing information from different geographical regions of the airway
- c. Explore relationships between both these and structural (CT), inflammatory outcome measures and electrophysiological measures.

2. Airway PD measurements:

- a. Develop the skills required to measure nasal PD
- b. Assist in the design and implementation of an improved protocol for measuring bronchial PD which is compliant with current regulations
- c. Obtain repeated measures of both nasal and bronchial PD in patients receiving a single dose of non-viral CFTR gene therapy
- d. Design a novel nasal PD technique to allow quantification of transfected cells.

2 Methodology

2.1 The Run-In study

The Run-In study was a longitudinal observational study involving patients with CF recruited in London and Edinburgh as part of the UKCFGTC clinical programme. As outlined in Chapter 1, section 1.5.1, it aimed to identify both the most appropriate outcome measures and the appropriate patients to take forward to a subsequent multidose trial of CF gene therapy.

2.1.1 Study personnel

The clinical team at each site included clinicians and research nurses at the University of Edinburgh, and clinicians, nurses and physiologists at the London site. Scientists and administrative staff were based at all three Consortium sites (the Oxford site is non-clinical). Recruitment was complete prior to the author joining the clinical programme; however the study has been in progress throughout the time course of this PhD. Over this period, the author's role has been predominantly in overseeing and refining the methodology of airway physiology tests, quality control and data analysis.

2.1.2 Study recruitment and consent

The Run-In study involved both adult and paediatric participants in Edinburgh (Western General Hospital and Royal Hospital for Sick Children) and London (Royal Brompton Hospital). Patients were recruited and enrolled in the study by the respective clinical teams at each site. Potential participants were enrolled if they gave informed consent (or assent with parental consent) and met inclusion criteria.

2.1.2.1 Inclusion Criteria

1. Cystic fibrosis confirmed by a combination of clinical symptoms/ signs and sweat test, genetic analysis or nasal potential difference measurements
2. Forced expiratory volume in the 1st second (FEV₁) > 40% predicted values*
3. Clinical stability at entry defined by:

- a. Not on any extra antibiotics (excluding routine, long-term treatments) for the previous 2 weeks
- b. No increase in symptoms such as change in sputum production/ colour, increased wheeze or breathlessness over the previous 2 weeks
- c. No change in regular respiratory treatments over the previous 4 weeks
- 4. Aged 10 years and above
- 5. Written informed consent from patient (aged 16 years and above) or parent if younger child; assent from child where appropriate
- 6. Permission to inform GP of participation in study

*Inclusion criteria for the Run-In study were based on spirometry reference equations as defined by Quanjer (61) (adults) or Rosenthal (62) (children).

2.1.2.2 Exclusion Criteria

- 1. Current participation in any other interventional clinical trial
- 2. Infection with *Burkholderia cepacia* complex organisms or MRSA unless local infection control guidelines could be adhered to
- 3. Awaiting/ referred for or post-lung transplant
- 4. Chronic day-time oxygen requirement
- 5. Previous spontaneous pneumothorax unless patient had had a subsequent pleurodesis
- 6. Recurrent severe haemoptysis
- 7. Current smoker
- 8. Significant co-morbidity including:
 - a. Moderate/ severe CF liver disease (varices or significant, sustained elevation of transaminases: (ALT/ AST>100 IU/l)
 - b. Significant renal impairment (serum creatinine > 150 µmol/l)
 - c. Significant coagulopathy
- 9. Receiving 2nd line immunosuppressant drugs such as methotrexate, cyclosporine, intravenous immunoglobulin preparations
- 10. Pregnant

There was no upper limit of FEV₁% predicted at recruitment. Patients were made aware at time of enrolment that the Run-In study was a precursor to a future

multidose trial of CF gene therapy but that the study itself was observational and participation in the Run-In study did not guarantee an opportunity for enrolment in the future multidose trial.

2.1.3 Study visits and clinical stability

Study visits were performed during periods of clinical stability at intervals initially aimed to be around 3-5 months. Clinical stability was confirmed by telephone call prior to the study visit date. Importantly, visits were rescheduled if the patient reported a change in usual symptoms or recent change in medications (over past 4 weeks), which resulted in many visit intervals being extended.

At the study visit, patients were first asked to complete a short questionnaire which included a series of questions designed to assess clinical stability (see Appendix 9.3). Questions related to change in symptoms of cough, tiredness, breathlessness and sputum (volume and character). For each question, response was on a 5 point scale from -2 to +2 with negative numbers representing deterioration in symptoms. Potential scores were between -14 and +14. Patients were required to have a symptom score above -4 to continue with the visit.

Tests were carried out at study visits as per trial protocol. Study visits took between two and three hours to complete per patient.

2.1.4 Study protocol

The Run-In study protocol was devised by the UKCFCTC Strategy Group. The Chief Investigator was Prof Eric Alton. A favourable ethical opinion was obtained in June 2007 (REC reference 07/Q0703/78). Patient enrolment commenced in January 2008. This thesis includes data from the first four study visits. It was at this time point that decisions were made regarding which outcome measures would be used and which participants would be eligible for the multidose trial. Following this decision, eligible patients were seen for visits 5 and 6 to ensure on-going longitudinal data collection prior to the start of the multidose trial. Those considered ineligible on the basis of disease severity (too severe or too mild) for inclusion in the multidose trial were offered the opportunity to attend a study 'exit' visit.

The Run-In study outcome measures investigated in this thesis are summarised in Table 2.1. A full list of outcome measures investigated in the Run-In study can be found in Appendix 9.1.

Test	Outcome measure(s)
Blood markers of infection and inflammation	Neutrophil count, CRP
Spirometry	FEV ₁ , FVC, FEV ₁ /FVC, MEF ₂₅₋₇₅
Multiple breath washout	LCI
Chest CT scan	Extent and severity of bronchiectasis, airway wall thickness, air trapping
Quality of life questionnaire (CFQ-R)	Respiratory domain

Table 2.1. Run-In study outcome measures described within this thesis

2.1.5 Data analysis

Data analysis described in this thesis is limited to the first 4 study visits. As described, this reflected the data on which decisions were made regarding appropriate assays and patients to take forward to the multidose trial. All data analysed from the Run-In study has been obtained from the official study database produced by the Royal Brompton Hospital Clinical Trials and Evaluation Unit (CTEU). The CTEU database was created in Microsoft Access. This database included a series of individual data tables summarising patient characteristics and outcomes.

Export of these Access data tables to Microsoft Excel was performed by the author. Patients were identified by their study number (RU001 to RU216) in each table. In order that serial measurements from individual patients could be aligned across study visits (to permit comparison between visits), additional ‘blank’ rows containing no data were inserted for any RU ‘number’ not actually representing an enrolled patient at visit 1. For visits 2-4, additional ‘blank’ rows were inserted as per visit 1, and also for any patients with no data for a particular visit due to prior

withdrawal from the study. These alignment processes resulted in the same patient being represented in the same row of data for each data table at each visit.

Visit interval independent results are summarised as box and whisker plots (range, 25th, 50th and 75th centile) for each outcome measure as a data summary of results from all patients with results from all 4 study visits. The coefficient of variation (%) between visits for an outcome at the 4 study visits was calculated by dividing the standard deviation of the result by its mean. This was performed for all patients with results for the outcome of interest at each visit unless stated otherwise in the text.

Cross-sectional correlations between computed tomography (CT) scores and other outcome measures were performed only on measurements made at the same Run-In visit (i.e. for a patient that had their CT scan at visit 2, comparison was made with the LCI measured at visit 2).

Relationships between continuous and rank ordered variables were explored with correlation coefficients (Pearson and Spearman rank). As this was an exploratory data set, the author undertook multiple comparisons between outcome measures; to compensate for this, the author has assigned statistical significance only to analyses with a p value of <0.0001.

Linear regression was used to investigate directional change over time. Using age in years as the x-axis variable, decimal age at each visit was plotted against different y-axis variables (e.g. lung clearance index) to calculate a 'best fit' regression equation for each patient. In the equation $y=mx+c$, where x was age in years, 'm' represented the change in variable per year. Comparison of the slope of change over time for an outcome of interest ('m' in the above equation) according to disease severity was performed with t-tests or Mann Whitney tests depending on normality of distribution. A p value <0.05 was considered significant.

Statistical analysis was performed using SPSS or GraphPad Prism software (version 4.03 GraphPad Software, Inc). Prism was used for graphical display of results. Statistical advice to the UKCFGTC has been provided by Dr Gordon Murray (GM), University of Edinburgh although formal statistical analysis of this dataset has not

yet been undertaken by his department. The analysis and results presentation described have been performed by the author with guidance from GM.

2.2 The single dose Pilot study

The Pilot study was a phase 1b clinical trial. It involved the administration of a single nasal and/or lung dose of non-viral CF gene therapy with outcome measures reflecting the study aims of investigating safety and functional efficacy.

2.2.1 *CFTR* gene therapy product pGM169/GL67A

2.2.1.1 Clinical plasmid

The *CFTR* plasmid DNA administered in the Pilot study was pGM169. Its nomenclature represents the number of plasmids investigated by the UKCFGTC in pre-clinical testing. Prior to the Pilot study, this plasmid had not been administered to man. However, it has been subject to an extensive process of pre-clinical testing within the UKCFGTC, leading to its selection for evaluation in clinical trial. To improve duration of *CFTR* expression, a humanised promoter rather than a viral (CMV) promoter (used in the Lancet study (35)) was inserted. In mouse studies, the human elongation factor 1 alpha promoter resulted in a sustained duration of expression in comparison to the CMV promoter (100).

The other major modification to the *CFTR* plasmid DNA was to remove all CpG dinucleotides from the plasmid DNA (100). Removal of these CpGs resulted in significant modulation of the post dosing inflammatory profile observed in mice (100). Plasmid pGM169 was superior in duration of expression and in the reduced likelihood of a post-dosing inflammatory response when compared with other plasmids hence it was selected to take forward for clinical evaluation.

The plasmid DNA used in the clinical programme was manufactured externally (VGX Immune Therapeutics, Texas, US). On site storage was maintained at -80 degrees C in pharmacy approved conditions until required for patient administration. Monitoring during storage on site included a locked temperature mapped freezer, within a controlled entry room.

2.2.1.2 Clinical liposomal vector

The UKCFGTC Pilot study used a non-viral gene transfer agent (vector) to increase the potential delivery of *CFTR* to the airway cells. This vector was a cationic liposome, GL67A (OctoPlus, Leiden, NL), identical to that administered during the Lancet study (35). It was found to be superior to other non-viral gene transfer agents (25KDa-branched polyethyleneimine (PEI), the cationic liposome GL67A and compacted DNA nanoparticles) in pre-clinical evaluation (101). GL67A is produced with three components: GL67, DOPE and DMPE-PEG5000. On site storage was undertaken as per clinical plasmid conditions, at -80 degrees C until required for patient administration. Despite the lipid agent itself being identical, the method of administration was different in the Pilot study reflecting knowledge gained from animal models and nebuliser droplet distribution. Pre-clinical optimisation of nebulisation including droplet size to increase the likelihood of the product reaching the small airways (thought to be the site of initiation of disease pathology in CF) has been undertaken by researchers within the Consortium prior to the Pilot study (102).

2.2.2 Trial protocol

The study design and protocol were devised by the Strategy Group of the UKCFGTC, which comprises senior clinicians and scientists from the three UK sites involved in the Consortium. The funding for the Pilot study was from the UK CF Trust and the trial Sponsor was Imperial College, London.

2.2.2.1 Study personnel

The clinical team at the London site was led by Dr Jane Davies (Reader, Honorary Consultant at the Royal Brompton Hospital, London and primary supervisor for this PhD). Other members include nurses, physiologists, trial co-ordinator and clinical research fellows (including the author of this thesis). The scientists at the London site were led Dr Uta Griesenbach, and the Chief Investigator was Professor Eric Alton. Sample processing (sputum and serum) was performed by laboratory technicians within the Department of Gene Therapy, Imperial College and at the Edinburgh Consortium site. Routine clinical laboratory tests (blood and sputum) were processed by the Royal Brompton Hospital laboratory services. Molecular

assays of *CFTR* were performed at the Oxford site of the Consortium, by the team of scientists led by Dr Steve Hyde.

The Pilot study took place throughout the first 2 ½ years of the author's PhD. The author had pivotal roles in the conduct of the trial: consenting patients, providing medical care for patients on bronchoscopy and dosing days, and performing airway potential difference (PD) measurements. During the study, the author also prepared summary progress reports for the UKCFGTC Strategy Group, Data Safety Monitoring Board (DSMB) and regulatory authorities. Post trial, the author's focus has been on data analysis. This has included all upper and lower airway PD measurements, as well as the safety and physiological outcomes.

2.2.3 Study enrolment

Participants were all registered patients at the Royal Brompton Hospital, London. Potentially suitable patients were identified according to baseline lung function. This was done by all members of the clinical team, but predominantly nursing and physiologists, often by introducing the study to patients at routine outpatient appointments where a Patient Information Sheet was given to interested persons. These patients were then followed up by telephone to determine if they were keen to participate and a study visit ('Initial visit') was arranged. Eligible patients were enrolled in the study if informed consent was obtained. Informed consent was taken by a member of medical staff (JD, author or TB). Compensation for loss of earnings and travel expenses for study visits were reimbursed.

At enrolment, patients were required to be clinically stable, and meet the trial inclusion and exclusion criteria as listed in the trial protocol as defined below.

The trial was open-label and was specifically designed to be iterative, amendments to protocol being made on the basis of responses seen in an earlier dosing group. Recruitment was opportunistic, based on subject availability. There were no specific attempts made to 'match' patients on the basis of clinical characteristics within dosing cohorts.

2.2.3.1 Inclusion criteria

1. Cystic fibrosis confirmed by sweat testing or genetic analysis
2. Males and females aged 16 years and above
3. Forced expiratory volume in the 1st second (FEV₁) > 60% predicted* value (or > 40% predicted* for patients receiving the nasal-only dose under Part B of the study)
4. Clinical stability at entry defined by:
 - a. Not on any additional antibiotics (excluding routine, long-term treatments) for the previous 2 weeks
 - b. No increase in symptoms such as change in sputum production/colour, increased wheeze or breathlessness over the previous 2 weeks
 - c. No change in regular respiratory treatments over the previous 4 weeks
5. Prepared to take effective contraceptive precautions for the duration of their participation in the study and for 3 months thereafter
6. If taking regular RhDNase were willing, and considered able by independent medical carers, to withhold treatment for 24 hours before and 24 hours after the gene therapy dose (nebulised doses only)
7. Written informed consent obtained
8. Permission to inform GP of participation in study

*Percent predicted values for lung function at recruitment were based on Quanjer (61) and Rosenthal (62) reference values for adults (≥ 18 years) and children (< 18 years) respectively. This reflected our current practice at that time of protocol design.

2.2.3.2 Exclusion criteria

1. Infection with *Burkholderia cepacia* complex organisms or MRSA (due to cross infection reasons)
2. Significant nasal pathology including polyps, clinically-significant rhinosinusitis, or recurrent severe epistaxis
3. Acute upper respiratory tract infection within the previous 2 weeks
4. Previous spontaneous pneumothorax without pleurodesis
5. Recurrent severe haemoptysis (Part A only)
6. Sensitivity, allergy or inability to take non-steroidal anti-inflammatory agents (Part B of study/ Ibuprofen arm only)
7. Current smoker (assessed by measuring cotinine level prior to study entry)
8. Significant co-morbidity including:
 - a. Moderate/severe CF liver disease (varices or significant, sustained elevation of transaminases: ALT/ AST>100 IU/l)
 - b. Significant renal impairment (serum creatinine > 150 umol/l)
 - c. Significant coagulopathy
9. Receiving 2nd line immunosuppressant drugs such as methotrexate, cyclosporine, intravenous immunoglobulin preparations
10. Pregnant or breastfeeding

2.2.4 Ethics, MHRA and study protocol amendments

The Pilot study was approved by GTAC (reference 140) and the MHRA and registered on the Clinical Trials database (EudraCT reference number 2007-004050-85). MHRA and GTAC approval for the protocol version (version 4) used at the start of the trial were granted in December 2008 and January 2009 respectively. Favourable approval from the Royal Brompton Hospital Research and Development department was obtained prior to enrolling the first patient.

The trial protocol was modified on several occasions during the course of the study. These modifications were designed to maximise the potential for obtaining appropriate information on product safety and dose response. Each protocol amendment was approved by the appropriate regulatory bodies. All patients

proceeding to the dosing study visit received a single dose of CF gene therapy (lung and/or nasal). No patients received placebo.

The Pilot study was conducted in two main sections, referred to as Pilot A and Pilot B.

2.2.4.1 Original trial design and protocol amendments (Pilot A)

Based on the previous Lancet trial (35), the proposed nebulised dose of gene therapy for this trial was 20 ml. The original protocol included an initial three patients receiving half this dose (10ml) nebulised to the lung and a 1ml nasal dose (a 'non-invasive cohort'). Following this, 12 patients would receive a 20ml lung dose and 2ml nasal dose with paired pre- and post- dose bronchoscopies. Post dose bronchoscopies were planned to occur on day 2 or day 14 post dosing.

However, the post dosing side effect profile seen in the 10ml lung dose patients led to the submission of a protocol amendment to a) include a 'non-invasive' cohort of patients (n=3) receiving 20ml to allow the natural history of any adverse effects to be documented without a bronchoscopy potentially complicating this and b) to move the day 2 bronchoscopy as at this time point, a flu-like response was still apparent; these procedures were therefore designed to be conducted on day 6 or 14 post dosing. Completion of the bronchoscopic cohort in late 2009 concluded this section of the Pilot study, referred to as 'part A'.

2.2.4.2 Amended trial design and protocol amendments (Pilot B)

Part B of the Pilot study (2010 - 2011) was initiated once it was clear that the standard delivery of 20 ml would not be sufficiently safe to proceed to multiple dosing. The two strategies employed broadly focused on a) trying to deliver the 20 ml dose more safely (slower nebulisation or with adjuvant anti-inflammatory therapies) and b) exploring the effects of lower doses. We also continued to investigate evidence of functional efficacy from nasal PD measurements and molecular assays. However, review of data by the independent Scientific Advisory Committee led to the recommendation that the focus of part B should be on safety outcomes and that the invasive bronchoscopic measures should be discontinued.

Patients in part B had lung and/ or nasal dosing. Protocol amendments were approved with removal of the bronchoscopic measures as above, and additionally chest CT scan and sputum cytokine assays were also taken out. CT was removed as there were no significant concerns raised following review of serial chest CT scans from patients in part A, and in an effort to reduce the time commitment required for study visits, sputum cytokine assays (and therefore sputum induction) were removed from the trial protocol.

2.2.5 Study visits

The number of study visits and tests carried out at each visit depended on the patient's dosing cohort. A summary of study visits according to dosing cohorts for both part A and B, including assays carried out at visits is provided in Appendix 9.2.

Post nebulised and/or nasal dosing, follow up was at pre-defined intervals until day 28. Within the constraints of the trial protocol, additional study visits were undertaken on the basis of adverse clinical status coinciding with the end of follow up or on-going evidence suggestive of functional efficacy (nasal PD changes) at day 28.

2.2.5.1 Pre-dose visits

Pre-gene therapy study visits were carried out during periods of clinical stability, as defined in the trial inclusion criteria. Pre dosing visits included the initial visit and a baseline visit (within 7 days of gene therapy dose), and additionally patients in the invasive cohort in Pilot A had a pre-dose bronchoscopy visit. All bronchoscopies were performed under general anaesthetic. Clinical stability on dosing day was confirmed by telephone call carried out by one of the research nurses prior to attendance and confirmed on the day of dosing according to symptom report.

Patients were advised to continue taking all regular medications throughout the study period, with the exception of RhDNase which was omitted for a period of 24 hours pre- to post gene therapy dose (due to a theoretical risk of degradation of the gene therapy product by the RhDNase).

2.2.5.2 Dosing visit

Clinical history and examination pre administration of lung dose was performed by a member of the medical team (author/JD/TB). Any patient not considered clinically stable on the basis of clinical history had their gene therapy postponed. A detailed description of procedures followed on dosing day, including administration of the gene therapy product, is given below.

2.2.5.3 Post dose visits

Post-dose visits were carried out at intervals determined by trial protocol, and according to patient dose cohort. The tests carried out at post dose visits depended on the dosing cohort and whether they were enrolled during Pilot part A or part B. Patients in the invasive (bronchoscopic) cohort had a single post dose bronchoscopy carried out.

2.2.6 Dosing: Preparation

Reconstitution and preparation of pGM169/GL67A was undertaken by the designated clinical trial pharmacists in a room within the Department of Gene Therapy, Imperial College suitable for purpose. Each vial of GL67A was hydrated with 5.25ml sterile water for injection and vortexed at 2500 rpm for 40 minutes (VWR-VX2500 vortexer (VWR International, PA, US)). After this period, the vial was checked to ensure the lipid was completely resolved and if not vortexing was repeated in 10 minute cycles (if still incomplete at 100 minutes the vial was discarded). In comparison for the plasmid (pGM169), once thawed, no reconstitution was required. To prepare the pGM169/GL67A complexes, a pneumatic mixing device was used (LMD2 pneumatic mixer, as described by Davies et al (103)). This device utilises a pneumatic piston, which in conjunction with a dual lumen syringe (Plas Pak, CT, US) and static mixer, achieves reproducible mixing of the pGM169 and GL67A components to produce the product ready for patient administration (5ml of the reconstituted lipid is added to one lumen, and 5ml of pGM169 to the other of the double lumen syringe). The gene therapy product (as a 10ml volume) was collected from the mixing device in a 20ml Sterilin container (Sterilin Ltd, UK).

Following preparation of the gene therapy product, it was delivered to the dosing cubicles in the Royal Brompton Hospital (RBH) in sterile universal containers in 10ml aliquots. Transportation was undertaken in accordance with a risk assessment carried out by Dr U Griesenbach: the product was transported within designated containers (shatter proof metal containers) containing absorbent material, with the person responsible for transporting the product carrying a 'spill kit' in case of accidental product release. It was an approximately 10 minute walk between the room used for product formulation and the area in the Royal Brompton Hospital used for dose administration. The reconstituted product had an expiry of 4 hours following preparation by the trial pharmacists.

On receipt of the product by study nursing or medical personnel within the RBH area designated for dose administration, a 10ml syringe was used to draw up the gene therapy into the nebuliser chamber or nasal spray bottle. This was done by study nursing or medical personnel.

2.2.6.1 Lung dose

For lung dosing, 5ml aliquots were inserted into a nebuliser chamber (AeroEclipse II Breath Actuated Nebuliser (BAN), Monaghan, NY, US). Further aliquots up to the total prescribed dose (in ml) were subsequently prepared.

2.2.6.2 Nasal dose

The nasal dose was administered using nasal spray bottles (GSK parts No. AR5989 30mL bottle/AR9488 30mL actuator). On receipt of the gene therapy product, the 10ml aliquot was transferred in its entirety to the nasal spray bottle, which was primed prior to use to ensure that the product would be delivered when required. This was done by inserting the spray bottle nozzle into an adapted 50ml Falcon tube (BD Biosciences, UK) which ensured that no product would enter the atmosphere following actuation. The Falcon tubes were part-filled with absorbent material so priming could occur into the tube within a contained space.

2.2.6.3 Weighing of nebuliser and nasal product

The AeroEclipse II nebuliser chamber and nasal spray bottle were weighed separately pre- and post filling, post-priming of the nasal spray, and post patient administration. For each 5ml nebuliser aliquot, this was undertaken when the chamber looked visibly empty and no ‘mist’ could be seen with patient use. For nasal dosing, post dose weighing was carried out when the designated number of nasal sprays had been administered. Weights were recorded to the nearest 0.1g on a worksheet, and a simple calculation was undertaken to determine the actual amount (assumption 1ml gene therapy product = 1g) delivered to the nose and with each 5ml nebulised aliquot. No ‘top up’ doses were given if the weight suggested under-dosing, either for nasal or lung delivery.

2.2.6.4 Risk assessment

All handling of the gene therapy product was made in accordance with the trial risk assessment, which included the use of personal protective equipment (FFP2 mask, apron and gloves) worn by staff entering the gene therapy delivery dosing cubicle, and also by patients if they needed to come out of the cubicle prior to the end of dosing and including a 30 minute period following the end of delivery. Following this period, no additional precautions regarding protective clothing or equipment were required either by patients or staff, other than existing hospital policy and procedure for segregation of patients with cystic fibrosis and infection control.

2.2.7 Dosing: Gene therapy administration

The nebulised dose was driven by pressurised air at 8 litres per minute using a breath activated nebuliser. An AeroEclipse II BAN kit with PARI filter set (041G0500, PARI Medical Ltd, Surrey UK) attached was used for each 5ml aliquot of gene therapy, with the mouthpiece from the PARI set used for dosing (Figure 2.1).



Figure 2.1. Lung dose nebuliser set up. The photograph shows an AeroEclipse II BAN nebuliser with PARI filter and mouthpiece, as used for lung dosing administration.

All dosing was carried out with the patients seated in small, individual cubicles. A glass window allowed continuous observation. Patients were instructed to breathe normally during nebulisation. Nebulisation was undertaken either at ‘standard delivery rate’ (3 minutes on the nebuliser and 2 minutes off to deliver each 5ml aliquot over approximately 25 minutes) or ‘slow delivery’ (1 minute on nebuliser and 4 minutes off, each 5ml aliquot delivered over a period of 75 - 150 minutes). Nebulisation was undertaken using the 5ml aliquots of gene therapy product, with the total number of aliquots administered dependent on intended dose as described above. All patients in Pilot part A received their nebulised gene therapy delivered at standard rate. Patients in Pilot B had either standard or slow rate delivery.

The nasal dose was administered by instructing the patient to deliver a single spray to each nostril once every five minutes during the nebuliser ‘off period’ in patients who were also being lung dosed; nursing staff recorded the number of sprays delivered. The number of sprays equal to 1ml or 2ml of gene therapy product had been previously determined prior to clinical trial by Consortium members at the Oxford site (1ml = 12 sprays, 2ml = 24 sprays). Nasal spray bottles were re-weighed post delivery of the appropriate number of sprays to determine actual dose delivered. Nasal dosing was administered according to the number of sprays rather than change in weight of nasal spray container.

The first three 10ml patients were dosed individually, 1 week apart. This was undertaken in view of this being a ‘first in man’ study to allow appropriate safety

monitoring. The first three patients in the 20ml nebulised (non-invasive) cohort were also dosed a week apart and admitted electively overnight for post-dosing monitoring. For subsequent dosing, the maximum number of patients treated on any day was 2, with considerations for patient safety and appropriate monitoring and staffing constraints paramount.

All patients in Pilot part A were recruited to receive nasal and lung dosing. Nasal and lung doses were administered simultaneously. Patients in Pilot B were recruited to lung, lung and nasal or a nasal only cohort. This was done according to patient preference, and also according to results of nasal PD testing (evidence of chloride secretion on pre-dosing tests excluded some patients from nasal dosing, see Chapter 5).

2.2.8 Clinical monitoring

On dosing day, and following confirmation of clinical stability, spirometry and baseline clinical observations were undertaken prior to dosing.

During dosing, patients could be directly observed through glass windows of the cubicles used for dose administration. They were asked to report any symptoms experienced during dosing to the nursing team. Oxygen saturation and heart rate were monitored every 30 minutes from the start of dosing.

Patients were observed for a minimum of 30 minutes in the dosing cubicles following completion of nasal and nebulised dose. A period of not less than 6 hours observation was then undertaken either in the Clinical Studies Unit/Biomedical Research Unit or as an inpatient at RBH. Spirometry, oxygen saturation, blood pressure, heart rate, respiratory rate and temperature were measured hourly post dosing for a period of 6 hours.

2.2.9 Adjunctive therapy

2.2.9.1 Planned

As per protocol, all patients (Pilot A and B) received a single 200 microgram dose of Salbutamol delivered via metered dose inhaler and large volume spacer 20 minutes

prior to dosing. This was administered because of an expected possible bronchoconstrictive effect of the hypotonic gene therapy product.

In selected lung dosing cohorts (Pilot part B only), planned adjuvant treatment was co-prescribed with the gene therapy. The aim of any adjuvant therapy was to assess its impact on the post dosing adverse event profile. Adjuvants investigated in the Pilot study included the oral drugs ibuprofen, prednisolone and paracetamol. The dosing strategies for these adjuvants is summarised in Table 2.2.

Planned adjuvant	All or selected patients only?	Dose and regime
Salbutamol	All	200 mcg (inhaled) via MDI and large volume spacer 20 minutes pre- nebulised dose
Ibuprofen	Selected only	400mg (oral) between 3 and 1 hour pre-dose with re-administration post dosing at intervals no less than 6 hours and no more than 3 doses in 24 hours
Prednisolone	Selected only	30mg (oral) daily from 24 hours pre- to 24 hours post nebulised dose (3 day total course)
Paracetamol	Selected only	1g (oral) at 2, 6, 12 and 18 hours post start of nebulised dose

Table 2.2. Adjuvant medication administered with lung gene therapy dose

Doses for planned adjuvant treatment with either ibuprofen or paracetamol were according to standard anti-inflammatory doses appropriate for adult patients. The dose selected for adjuvant prednisolone therapy was undertaken following consultation with a senior researcher with expertise in corticosteroids and relationship with inflammatory processes in chronic lung disease external to the trial employed within our institution (Professor P Barnes, National Heart and Lung Institute, Imperial College).

2.2.9.2 Symptomatic

Additional medication was permitted post dosing for symptomatic relief, including bronchodilators and simple analgesia.

In addition to any planned adjuvant treatment for patients in part B of the study; patients could also receive symptomatic treatment with ibuprofen and

paracetamol. This was prescribed on an ‘as-required’ basis according to patient request and symptoms (as long as drug or dose interval was not contraindicated by existing adjuvants).

2.2.10 *Adverse events and safety reporting*

Adverse events during the Run-In and Pilot study were collected on all patients from trial enrolment. Methods of identifying adverse events included patient report, patient symptom diary, together with review of clinical laboratory blood results. Serious adverse events (SAEs) and suspected unexpected serious adverse reactions (SUSARs) were reported to the trial Sponsor and relevant bodies.

2.3 Outcome measures: Airway physiology

This section includes methodology for airway physiology outcomes in the Run-In and Pilot studies. This includes extensive development work undertaken by the author to improve data quality and analysis for spirometry and LCI outcomes. Where results are presented to illustrate these quality and analysis issues, they are from the Run-In study Visit 1.

2.3.1 *Spirometry*

For both Run-In and Pilot studies, spirometry was performed on all subjects at trial enrolment and subsequent study visits according to ATS/ERS standards (51). An exception to this was in the six hour post gene therapy dose monitoring period in the Pilot study. During this acute monitoring period, repeated measurements were made hourly for a period of 6 hours as per protocol, however patients often found this exacerbating of any respiratory symptoms post dose and therefore we considered it acceptable that patients perform only a single blow.

2.3.1.1 Equipment

Spirometry was performed at each visit using an EasyOne[®] spirometer (ndd Medical Technologies, Zurich, Switzerland), with a single disposable mouthpiece and filter (spirette[®], ndd Medical Technologies as above). The EasyOne[®] is a portable, battery operated spirometer which uses ultrasonic flow sensors.

2.3.1.2 Calibration

A calibration check was performed on the day of test with the EasyOne® using a 3 litre syringe, with acceptable results if $\pm 3\%$. Due to the ultrasonic flow detection system used by the EasyOne®, it does not require calibration by the researcher (this is because the ultrasonic transducers are external to the spirette® and therefore changing the spirette® does not alter the transducers, which also function independent of temperature, atmospheric pressure and humidity).

2.3.1.3 Data collection

Spirometry tests were performed by study personnel (physiologists, nursing staff, clinical research fellows) at London and Edinburgh sites. They were performed with the patient seated using an open circuit method. Patients were asked to wear a nose clip for testing. A minimum of three forced expiratory manoeuvres from maximal inspiration were undertaken by subjects at each visit. EasyOne® spirometers were prepared by selecting 'best value' and 'Automated QC' in the test settings menu. Patient data was entered with study number as identifier, date of birth and height. 'FVC(expiration)' was selected in test selection menu, prior to insertion of the spirette®. Baseline flow setting was performed with occlusion of the base of the spirette®, following which a 'blast out' prompt was displayed on the EasyOne® screen. With patients holding the EasyOne® spirometer, they were instructed to place the mouthpiece into their mouth and form a complete seal around it. Patients then performed a forced sustained expiratory manoeuvre from a maximal inspiration. Verbal encouragement was given to subjects during testing. Visual inspection of the flow volume curve on the EasyOne® screen after each blow was made by the researcher carrying out the test. Three forced expiratory manoeuvres were performed, with measurement of FEV₁, FVC, and expiratory flows.

The raw flow and volume measurements captured for each EasyOne® test included FEV₁, FVC, MEF₅₀, MEF₂₅, MEF₇₅, MEF₂₅₋₇₅, and FEV₆. Selection of FEV₁, FVC and MEF₂₅₋₇₅ was performed according to ERS/ATS standards (51). Within manoeuvre, each blow had to demonstrate an absence of artefacts (e.g. cough during first second) and satisfactory duration of expiration (at least 6 seconds (3 for children)).

Between tests, across 3 acceptable tests the two largest volumes for FEV₁ and FVC had to be within 150 ml of each other (of 100ml if FVC < 1 litre).

The highest acceptable FEV₁ and FVC did not necessarily have to originate from the same blow (51). MEF₂₅₋₇₅ was taken from an acceptable blow with the highest sum of (FEV₁+FVC). Where two blows had identical sums of (FEV₁+FVC) but different MEF₂₅₋₇₅ values, the blow with the highest FEV₁ contributed to the final data value for that visit.

Parameters from all three acceptable forced expiratory manoeuvres were recorded in the patient's case record form (CRF), with selection of the final result for FEV₁, FVC and MEF₂₅₋₇₅ at each visit according to ERS/ATS standards. In both Run-In and Pilot studies, raw values were written into the CRF from the EasyOne[®] screen or exported directly to a computer via USB connection. The former was only relied on where the electronic data was unobtainable. Connection of the EasyOne[®] to a computer by placing it in the EasyOne[®] cradle, and using EasyWare[™] software resulted in an automatic upload of data. Print out of the collected data was kept for source data reference. The electronic data was imported from the EasyOne[®] software into Excel spreadsheets.

2.3.1.4 Original spirometry reference source

Raw values for FEV₁ and FVC (litres) and MEF₂₅₋₇₅ (litres/minute) were initially converted to % predicted values according to Rosenthal (<18 years) and Quanjer (≥18 years) reference equations (61;62). The reference equations used to determine the expected volumes for subjects according to these sources for FEV₁ are described in Table 2.3.

Reference source	Relevant group	Regression equation for % predicted FEV ₁
<18 years: Rosenthal et al (62)	M <162.6 cm	(0.03425*standing height in cm) -2.78
	M >162.5 cm	(0.0521*standing height in cm) -5.108
	F <152.6 cm	(0.03316*standing height in cm) -2.734
	F >152.5 cm	(0.04112*standing height in cm) -3.680
≥18 years: Quanjer et al (61)	M 18-70 years, height 1.55-1.95 ^{\$}	(4.3*standing height in m)-(0.029*age yrs)-2.49
	F 18-70 years, height 1.45-1.80 ^{\$}	(3.95*standing height in m)-(0.025*age yrs)-2.60

Table 2.3. Linear regression equations for calculation of FEV₁ % predicted as used at start of Run-In and Pilot studies. M=male, F=female. ^{\$}25 substituted for subjects between 18-25 years. ‘Relevant group’ refers to height and gender. Height is inserted into the relevant regression equation in centimetres for Rosenthal and metres for Quanjer reference equations respectively.

2.3.1.5 New spirometry reference source

During the time course of the Run-In study, published reference look-up tables spanning paediatric and adult age ranges became available (64). An Excel add-in downloaded from www.growinglungs.org.uk (and currently accessible from www.lungfunction.org) allowed conversion of raw spirometry values to % predicted, z scores for FEV₁, FVC and MEF₂₅₋₇₅ according to the reference source described by Stanojevic et al (64). In Excel, data were analysed with columns designated for the variable of interest (e.g. FEV₁), aligned with patient gender, decimal age and height (to nearest cm in Run-In and to nearest mm in Pilot study) at study visit. All lung volumes (litres) and flows (litres/minute) were recorded to 2 decimal places in the CRF. Using the macro in the add-in, conversion to % predicted values and z scores was performed for FEV₁, FVC and MEF₂₅₋₇₅. Height was measured on a single occasion at the start of the Pilot study as all participants were adults, therefore all % predicted and z scores were obtained from the same height measurement. As the Run-In study included children and adolescents, height was measured at each study visits for all participants.

2.3.1.6 Comparison between percent predicted values for FEV₁ obtained with original and new reference source

Comparisons between the percent predicted values obtained with the original and new reference source were investigated. Using the Run-In study data set at visit 1 as an example, the differences in absolute FEV₁ % predicted values according to original and new reference sources for the same raw data are plotted in Figure 2.2.

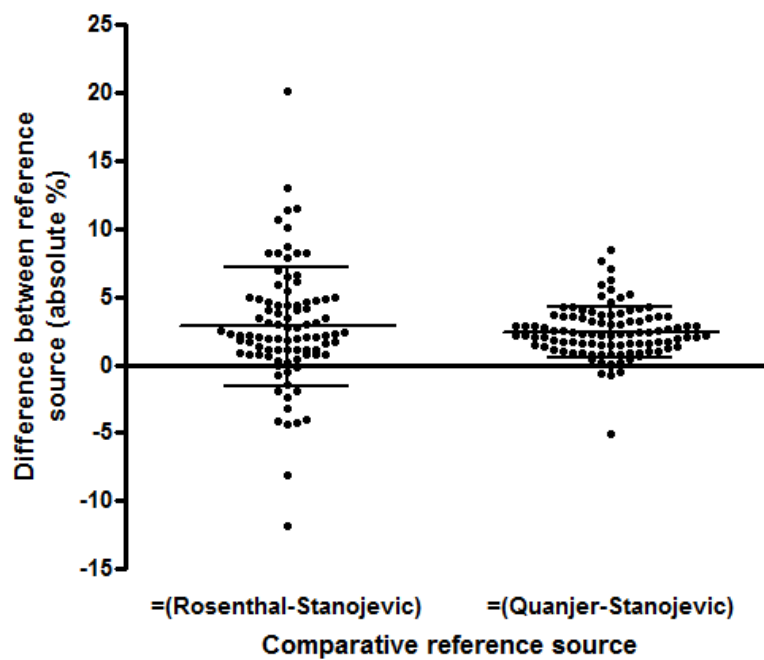


Figure 2.2. Difference in absolute FEV₁ % predicted values for identical raw data according to reference source at Run-In visit 1. Data on the left is the difference between Rosenthal and Stanojevic reference values for all patients aged <18 years at visit 1 with each patient represented by a single point. Data on the right is the difference between Quanjer and Stanojevic reference values for all patients aged ≥18 years at visit 1, with each patient represented as before. Lines represent mean and standard deviation.

Rosenthal and Quanjer FEV₁ % predicted values were both significantly higher than the Stanojevic values calculated with the same raw data. For all patients at visit 1, Rosenthal and Quanjer equations resulted in FEV₁ % predicted values on average 2.6% greater in absolute % predicted terms (paired t-test $p < 0.0001$). Divided according to paediatric and adult age ranges, the mean difference was 2.8% (95%CI 1.9-3.7) for children with Rosenthal equations ($p < 0.0001$), and 2.4% (95%CI 2.1-2.8) for adults using Quanjer equations ($p < 0.0001$). This difference between equations was present across the range of severity in disease in the Run-In study, illustrated by the Bland Altman (104) plot in Figure 2.3 with mean bias and 95% limits of agreement shown.

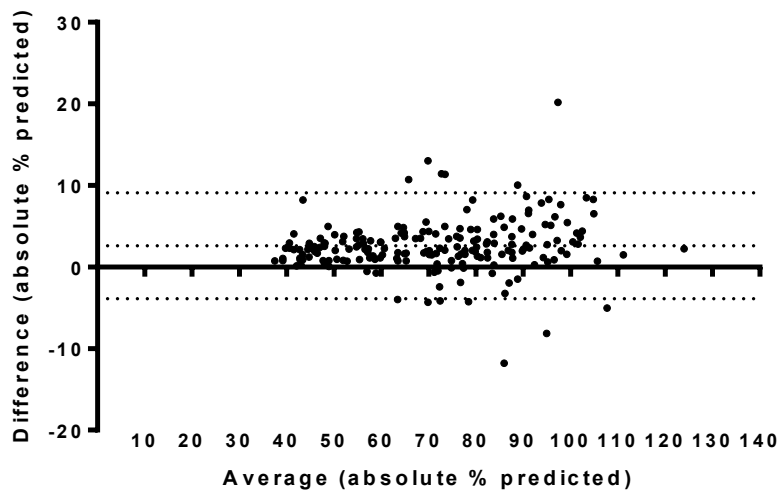


Figure 2.3A

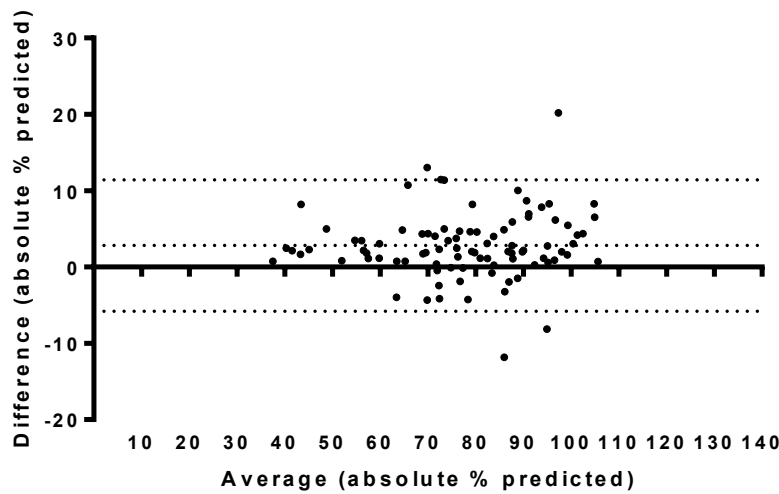


Figure 2.3B

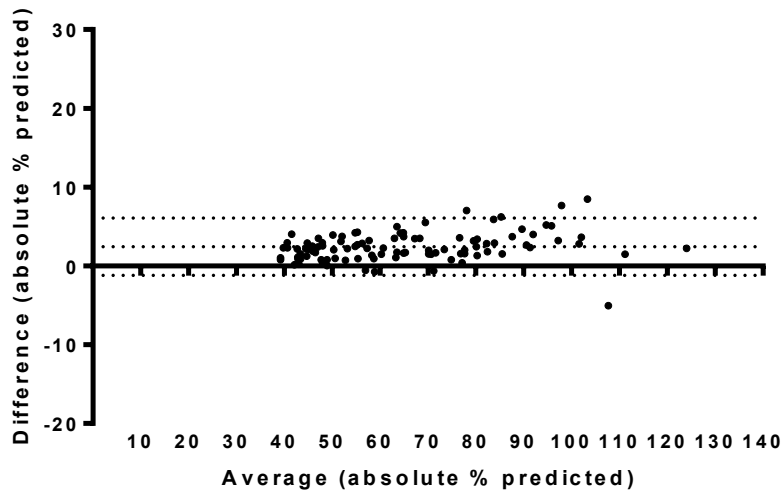


Figure 2.3C

Figure 2.3. Bland Altman plot for absolute FEV₁% according to reference source. FEV₁% predicted values are according to Rosenthal (<18 years) or Quanjer (≥18 years) compared to Stanojevic values for all patients (n=191) at Run-In visit 1 in (graph A), patients <18 years at visit 1 (graph B) and patients ≥18 years at visit 1 (graph C). The mean difference (bias) and 95% limits of agreement between Rosenthal and/or Quanjer vs. Stanojevic % predicted values are represented by dotted lines.

For patients aged <18 years at visit 1, the greatest difference between reference sources was seen for female participants with height <152.6cm at visit 1 (mean difference between Rosenthal and Stanojevic % predicted value = 6.2, paired t-test $p<0.0001$) although the number of patients in this bracket was small ($n=13$). Large discrepancies were also noticed between males aged <18 at visit 1 with height <162.5cm ($n=31$ patients, mean difference between Rosenthal and Stanojevic % predicted value = 3.7, paired t-test $p<0.0001$).

A plot of FEV₁ % predicted values calculated with the Rosenthal and Quanjer equations at visit 1 in comparison to the Stanojevic z score for FEV₁ at visit 1 is shown in Figure 2.4A. Similarly the FEV₁ % predicted values from the Stanojevic reference source in comparison to z scores are plotted in Figure 2.4B; as expected a tighter correlation was seen due to these measures being derived from the same data source. However, where discrepancies exist, this comparison demonstrates the additional benefit of using z scores to describe results as they take account of the expected normal distribution according to age, gender and height of the participant.

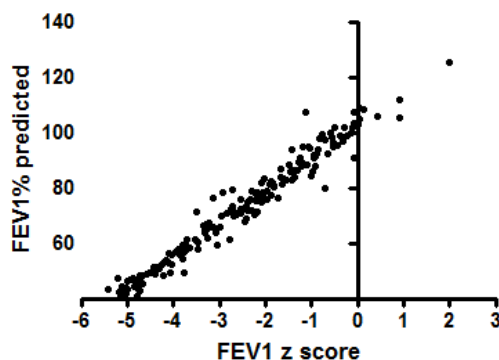


Figure 2.4A

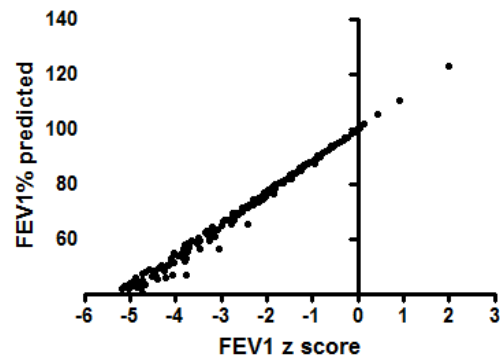


Figure 2.4B

Figure 2.4. Comparison of FEV₁ % predicted and z scores. In graph A, visit 1 FEV₁ % predicted values according to original reference (Rosenthal <18 years or Quanjer ≥18 years) are plotted against Stanojevic z scores. In graph B, % predicted and z scores from the Stanojevic source are compared. Each patient is represented by a single data point on graph A and graph B.

2.3.1.7 Patients crossing their 18th birthday during the longitudinal study

In view of the longitudinal nature of data collection in the Run-In study and the study population including both child and adult participants, it was expected that several patients would transition across their 18th birthday during the period of data collection. This occurred for 10 patients during the study, therefore crossing from Rosenthal to Quanjer reference sources at some point between visits 1 and 4 with the original reference source. This transition was often met with a steep step down in % predicted FEV₁ even in the presence of consistency with height and raw FEV₁ volumes.

For these 10 patients the FEV₁% predicted values at the study visit either side of their 18th birthday according to Rosenthal (<18) and Quanjer (after 18th birthday) or Stanojevic (all ages) are shown in Table 2.4.

Patient	Last visit before 18 th birthday: Rosenthal reference	1 st visit after 18 th birthday: Quanjer reference	Difference in absolute % predicted value(A)	Last visit before 18 th birthday: Stanojevic reference	1 st visit after 18 th birthday: Stanojevic reference	Difference in absolute % predicted value (B)
1	109.0	98.0	10.9	100.7	97.3	3.4
2	88.4	68.9	19.5	83.5	70.2	13.3
3	102.6	94.3	8.3	96.0	91.9	4.1
4	109.7	90.4	19.3	101.5	89.4	12.1
5	86.0	81.5	4.4	79.4	80.9	-1.5
6	79.0	72.5	6.4	74.3	72.3	1.9
7	75.9	66.3	9.6	71.4	65.8	5.6
8	35.8	30.7	5.1	35.6	30.4	5.2
9	74.3	65.0	9.4	61.4	64.3	-2.8
10	45.6	41.8	3.8	43.0	43.4	-0.4

Table 2.4. FEV₁% predicted values for the n=10 patients that crossed their 18th birthday during the first 4 visits of the Run-In study. Data shown are the % predicted values for the last visit before their 18th birthday, first visit afterwards, and the difference in absolute % predicted values between these visits according to Rosenthal to Quanjer (A) or Stanojevic (all ages) reference source (B).

The mean absolute drop in FEV₁ % predicted values across the study visits spanning a patient's 18th birthday was 9.7 absolute % points with Rosenthal to Quanjer equations in comparison to 4.1 absolute % points when Stanojevic reference values were used.

The slope ('m') representing annual decline in absolute FEV₁ % predicted values were calculated using the linear regression equation $y=mx+c$ for FEV₁% predicted values at each visit, taking into account the interval between study visits. Using this approach, the annual decline in FEV₁ % predicted values was significantly greater with Rosenthal/Quanjer than with Stanojevic reference source (paired t-test $p=0.0007$) (Figure 2.5), largely due to an artefactual drop when switching from Rosenthal to Quanjer values at age 18.

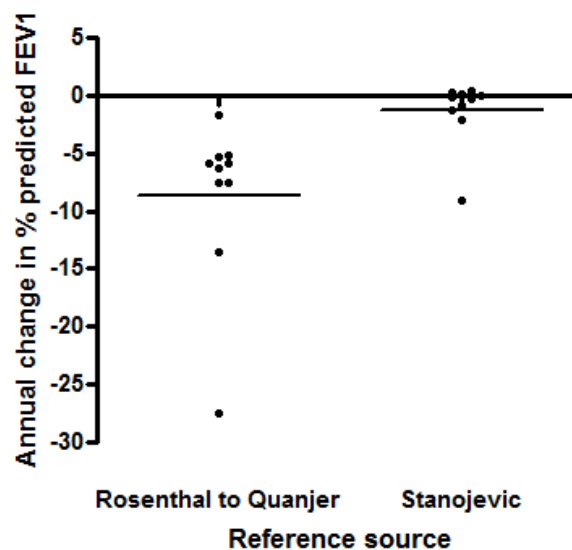


Figure 2.5. Annual change in absolute % predicted FEV₁ for the 10 patients transitioning across their 18th birthday during the Run-In study. Plotted according to linear regression analysis. Annual change is plotted according to the separate child and adult reference source (Rosenthal to Quanjer) or Stanojevic reference source alone.

2.3.1.8 Summary

In view of the above results, which are particularly relevant for any longitudinal study involving paediatric and adult participants, all raw values from Run-In and Pilot tests were converted to percent predicted values and z scores using the Stanojevic 'All Ages' spirometry reference source (64). All final spirometry percent predicted values and z scores reported in the Run-In and Pilot study results chapters (Chapter 3 and 4 respectively) were those obtained using the 'All Ages' Stanojevic reference source (64).

2.3.2 Lung clearance index

This methodology section applies to both the Run-In and Pilot studies. The author was involved in performing many of the tests, although the majority of data collected from the London centre was by study physiologists. For Run-In study patients in Edinburgh the test was performed largely by the Clinical Research Fellows. The author has led in identifying and responding to quality control issues, and been extensively involved in optimising analysis parameters and performing data analysis.

2.3.2.1 Equipment

Lung clearance index (LCI) was measured using a multiple breath washout (MBW) technique with sulphur hexafluoride as the inert tracer gas. MBW was undertaken using an Innocor (Innovision, DK) machine adapted for testing according to the methodology described by Horsley et al (48). The Innocor was originally developed for measurement of cardiovascular parameters by inert gas rebreathing. 0.2% SF₆ in air was used as the tracer gas (2000 ppb SF₆/air 200bar, cylinder BOC 159499-L-PC).

2.3.2.2 Calibration

The Innocor machine was switched on at least 20 minutes prior to testing to allow the photoacoustic gas analyser to warm up. Flow gas delay calibration was performed prior to the first test each day and repeated after 4 hours if further testing was planned. This was performed with a filter (details below) present

within the calibration set up. The researcher performing the MBW test then performed a series of 11 manoeuvres involving slow expiration followed by rapid inspiration. Each breath was separated by a few seconds. Acceptability criteria included the variation in flow gas delay between breaths being not more than ± 20 milliseconds. Flowmeter calibration was also performed at intervals as above. This was performed with a filter in situ using a 3 litre calibration syringe. Repetitions of filling and emptying the syringe (at different flow rates - slow x 2, medium x 2, fast 2) were performed; differences were required to be within $\pm 2\%$. Measured syringe volumes were also required to be within $\pm 2\%$ accuracy.

2.3.2.3 Performing the test

The MBW test was performed at each clinical site (RBH, RHSC Edinburgh and Western General Hospital). The test was performed with subjects at rest, distracted using a DVD, seated with a nose clip occluding the nasal passages. Tidal breathing was required throughout.

2.3.2.4 Wash-in phase

A single use disposable filter (Bacterial/viral filter Air Safety Ltd, Adult: 4000/01 Paediatric: 9070/01) was sited between the mouthpiece and flowmeter (Hans Rudolph Adult: 4700A, Paediatric: 4719). Subjects inhaled the gas mixture which was delivered via a T-piece flow past (Figure 2.6).

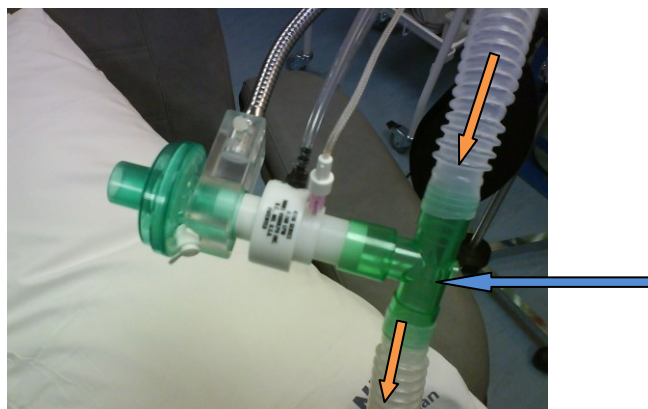


Figure 2.6. Illustration of T-piece flow past for LCI measurement. A mouthpiece is connected to the filter (on left) prior to patient testing. The T-piece is illustrated by the blue arrow. The orange arrows represent the flow of SF₆ in air in the tubing.

This was delivered at a rate of 10-15 litres/minute to ensure that expired gas was not re-inspired. With progression of the test, the difference in inspired and expired SF_6 was reduced as expired concentrations equilibrated to those in the cylinder. As the Innocor determines the difference in concentration between the sampling inlet and an internal reference within the machine, measures were taken to minimise the presence of SF_6 around the machine. These included opening doors and window (where possible) and use of a fan. The researcher was able to identify when wash-in was complete by noting a difference of $\leq 0.003\%$ in SF_6 concentration for three breaths on the screen of the Innocor machine (this concentration difference was deliberately more strict than the criteria for acceptable wash-in at the data analysis stage, to be certain of reaching adequate wash-in prior to the washout phase).

2.3.2.5 Disconnection

When equilibrium between inspired and expired SF_6 was reached, the ‘wash-in’ phase was complete and the subject was switched to breathing room air. The T-piece (and therefore gas supply) was disconnected when the patient was in expiration with rapid removal of T-piece from the vicinity of the patient to prevent any additional further inspiration of test gas. Tidal breathing of room air continued until the test gas had been ‘washed out’ of the lungs. This point was identified by the researcher on the Innocor screen as the point at which the maximal SF_6 concentration was of $\leq 0.003\%$ for 3 consecutive breaths. An example trace illustrating the wash-in and wash-out phases is shown in Figure 2.7.

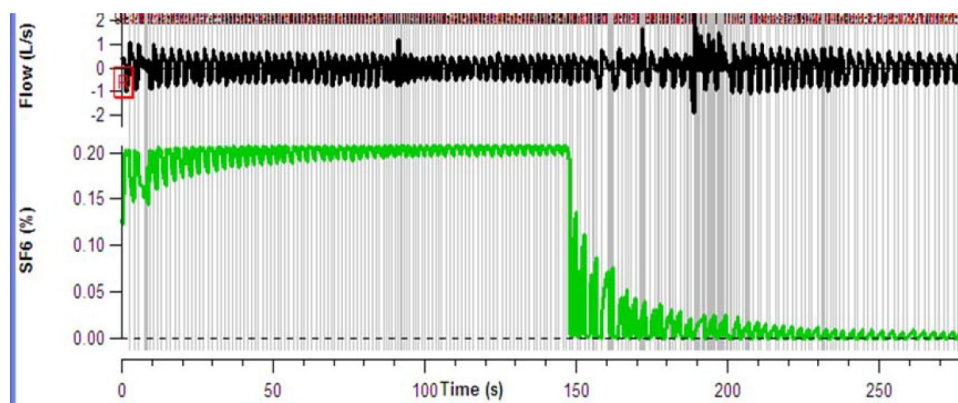


Figure 2.7. Multiple breath washout trace illustrating wash-in and wash-out phases of the test.

On each test date occasion, the procedures above were repeated until 3 tests had been performed.

2.3.2.6 Data export

Raw data from each test (1 file per washout) were identified by the format: [patient ID]-[Date]-[Time].ino and Innocor data files (.txt file extension). They were retrieved from the Innocor and saved on a network drive at each clinical site.

2.3.2.7 Data analysis

The process of data analysis for LCI evolved considerably during the Run-In and part A of the Pilot study in 2009-2010, due in large part to the work of the author.

2.3.2.8 Original analysis

At the start of the Run-In and Pilot studies, LCI tests were analysed using software written by Per Gustafsson (Queen Silvia Children's Hospital, Gothenburg, Sweden) (no commercially supported software for the standard Innocor for LCI analysis is available). During both Run-In and Pilot studies, a number of technical problems were detected during data analysis. A key trigger for detailed investigation of the LCI results occurred as a result of two analysers sensing SF₆ gas concentration reaching a different result for the same test, in combination with an unexpectedly high inter-visit variability for some patients in the Run-In study. Concerns with the quality and consistency of data analysis led to tests being re-analysed by a five member team across the two clinical Consortium sites. Between London (the author and VC (respiratory physiologist)) and Edinburgh (Clinical research fellows NB, HS and PR) sites, just under 2000 LCIs were reanalysed using the software written by Per Gustafsson. However, during this reanalysis process and on-going data collection it was also noted that not all traces had an end-inspiratory concentration of SF₆ of zero at completion of a washout. This had great potential to influence the LCI result, but it was not possible to adequately address this within the original software used. In response, NB wrote a new software programme ('SimpleWashout') for analysing the LCI data, and in doing so we found that there was a small but likely important 'offset' between the start and end ambient SF₆ concentration. Further investigation by NB with Innovision (Odense,

DK) confirmed that this was due to the methodology employed by the Innocor machine in determining gas concentrations relative to ambient (air) reference rather than from absolute values, thus it was not always possible to achieve a 'zero' SF₆ end expired gas concentration at the end of the washout.

The problem with offsets was particularly relevant for the London site. During the period of data collection described, the clinical area for LCI tests within RBH changed on 2 occasions in response to the building of a new Clinical Research Facility within the hospital. Problems with room ventilation in the interim temporary accommodation were considered likely contributory to the offsets.

2.3.2.9 Revised analysis for final LCI results

The final values for LCI described for the Run-In and Pilot study in Chapters 3 and 4 have been calculated using the SimpleWashout software (NB, University of Edinburgh). This programme was written using the IGOR Pro (WaveMetrics Inc) data analysis platform. The following parameters were entered into the SimpleWashout software window for calculation of the MBW parameters: Analyser, study, patient, visit, test, Flow gas delay and Height (m) - this allowed calculation of the minimum termination volume ($2.9 \times \text{height(cm)} - 215$, in mls) for aiding identification of the end of a washout.

Data files were imported into the SimpleWashout software by first converting the .ino files to .txt files containing flow, SF₆ and CO₂ data using the accompanying 'Innofileconverter'. The SimpleWashout software was written to minimise input required on the part of the researcher (to minimise operator impact on LCI result). Therefore only minimal input was required, however these were important and included:

1. Setting the disconnection point and grading the quality of disconnection (e.g. ensuring it had not taken place during inspiration).
2. Ensuring that the position of the cursors for C_{max} (concentration of SF₆ at end of last inspiration before disconnection) and C_{init} (average concentration of SF₆ at the end of the last expiration before disconnection) were correct.

3. Verifying that the first breath and subsequent breaths had been correctly identified
4. Checking the washout termination status
5. Checking the SF₆ offsets
6. Checking Expired: Inspired gas volume ratio

2.3.2.10 Calculating LCI

Fall in concentration of SF₆ in expired gas during tidal breathing of air, in combination with measurement of flow (and thus volume determination) is used to calculate the LCI. This is defined as the number of lung turnovers of functional residual capacity (FRC) taken to reduce the expired concentration of SF₆ to 1/40th of the 'wash-in' value.

The following criteria were used to determine acceptable quality control for tests used:

- Variability in concentration of test gas at wash-in was <0.008%
- Minimum difference between MBW tests was acceptable if FRC within 11% (test excluded if >11% different).
- Minimum difference between MBW tests within 21% for LCI (tests were excluded if >21% different)
- Satisfactory disconnection and termination of MBW test
- At least 1 other acceptable test on same test date (to allow calculation of average LCI and FRC).
- Expiratory: Inspiratory volume ratio <0.8 or >1.3

Repeat reanalysis of all Run-In LCI traces was performed with the NB SimpleWashout software, taking the above criteria into account to identify acceptable tests. Using this new software, 1961 Run-In traces were analysed between NB, GD, VC and PR, with HS performing the revised analysis of Pilot LCIs. Each Run-In and Pilot LCI test was then also subsequently been reanalysed by a second (different) operator to apply the offset and address any outstanding issues

regarding acceptability of data for inclusion in the study datasets. Inter-operator agreement was not formally investigated, as on the very rare occasion (since the process was largely automated) where a second analyser disagreed with the first, a consensus review process involving remaining members of the group determined the final outcome.

The 'SimpleWashout' software allowed an 'offset' correction for any SF₆ gas present in the room at the beginning of subsequent tests. The difference in LCI between tests performed at the London site with and without the Offset correction is displayed in Figure 2.8. The mean difference was negligible (0.02 LCI units). However, for individual patients the difference in LCI using the offset correction was more substantial. As extreme examples, one patient's value changed from 9.26 to 8.30 as a result of applying the offset and another from 12.18 to 13.58. These differences exceed changes that have been reported in response to clinically established interventions (69;70).

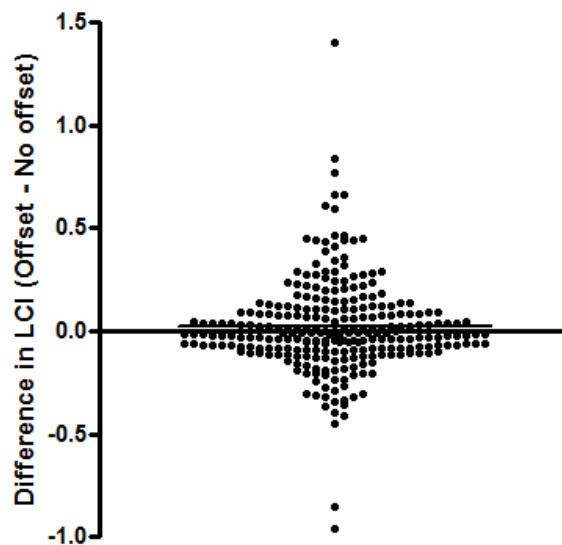


Figure 2.8. Influence of Offset correction on LCI result. Graph shows the difference in LCI as a result of application of Offset correction for all Run-In LCI tests performed at the London site, with each test represented by a single data point.

The results from the 'Offset corrected' LCIs are those contributing to the final values for Run-In and Pilot LCI and FRC results. Three tests were performed at each

Run-In or Pilot study visit, with the mean of 2 or more acceptable tests contributing to final LCI result.

2.3.2.11 LCI data loss

The process of data analysis described above led to the exclusion of tests not meeting the quality control criteria. These largely included failure to appropriately ‘wash-in’ with test gas prior to disconnection, or tests terminating prematurely. This led to an incomplete LCI data set for a proportion of patients (Table 2.5).

	All sites N (%)	London N (%)	Edinburgh N (%)
Total number of expected LCIs from all Run-In study visits	678	378	300
Total number of acceptable LCI results	568 (84%)	277 (73%)	291 (97%)
Total number of excluded tests/absent LCI results	110 (16%)	101 (27%)	9 (3%)

Table 2.5. LCI data loss following SimpleWashout analysis and quality control review, according to Run-In study sites.

The extent of data loss at the London site was disproportionate to loss of LCI test results performed in Edinburgh. It is likely this reflected factors beyond environmental influences in test gas measurement including an inadequate appreciation of the importance of adhering to the wash-in and washout criteria; which has since been resolved through further training and a more detailed standard operating procedure. However, also of note was a finding that tests across both sites were more likely to be excluded in patients with a visit 1 z score of <-2.5 (the Run-In study population median FEV₁ z score at visit 1). These patients accounted for 6/9 (67%) of the excluded tests from Edinburgh and 68/110 (62%) excluded tests from London; the author speculates this may in part be due to a longer period of time required to achieve adequate wash-in or washout.

The exclusion of tests according to the quality control rules described above has ensured quality and confidence in the final dataset on which multiple comparisons with other Run-In study outcomes were made (Chapter 3). Due to parallel timelines of the Pilot study and visits 3 and 4 of the Run-In study, these problems with data

loss at the London site particularly affected the patients in Pilot part A, most notably the 20ml lung dose cohort (Chapter 4).

2.3.3 Gas transfer

Transfer factor of the lung for carbon monoxide (TLCO) (referred to as 'gas transfer' throughout this thesis) was measured for all patients receiving a lung dose of gene therapy in the Pilot study using a single breath technique. Gas transfer was not part of the Run-In study protocol.

Irrespective of Pilot study dose cohort, testing was carried out on a single date pre-dose, and on days 2, 14 and 28 post gene therapy. In Pilot part A, where the post dosing measurement on day 14 coincided with post dosing bronchoscopy gas transfer was always measured pre- bronchoscopy.

Pilot study measurements were all performed using a Jaeger® Masterscreen PFTPro (CareFusion, Germany) by the trial physiologists. Volume and gas calibrations were performed prior to testing. Subjects were seated wearing a nose clip, and measurement undertaken using the diffusion single breath protocol. Patients were asked to exhale to residual volume, then take a maximal inspiration (required to be >90% vital capacity) of the test gas (medical air with 0.28% carbon monoxide and 9% helium) prior to a breath hold of 10 seconds duration. Tests where the subject expired during this period (detected by a near patient shutter >3Kpa) were invalid. Following breath hold, the subjects were asked to breathe out smoothly.

This procedure was repeated at least twice on each test day (minimal interval between tests 4 minutes). Duplicate estimates of TLCO had to be within 10% or 1.0mmol.min⁻¹.kPa⁻¹ (whichever greater) to meet requirements for test acceptability. At least 2 technically acceptable tests were obtained and the mean values used for test results. Up to 10 measurements could be performed to obtain these reproducible values. The following values were recorded for each test: Single breath TLCO, Alveolar volume (VA), Transfer coefficient (KCO)(=TLCO/AV) and Inspiratory vital capacity (IVC). TLCO and KCO were corrected for haemoglobin (Hb); TLCOc and KCOc respectively; either from Hb on day/nearest day to test date using the following formulas (105):

Males: $TLCOc = TLCO * (10.22 + [Hb]) / (1.7 * [Hb])$

Females: $TLCOc = TLCO * (9.38 + [Hb]) / (1.7 * [Hb])$

The gas transfer results from the Pilot study were analysed as change in absolute values and % change from baseline at each post gene therapy test date.

2.4 Outcome measures: Clinical, laboratory, structural and patient reported

2.4.1 Physical examination and pulse oximetry

2.4.1.1 Run-In

Height was recorded at each study visit for both adults and children using a Leicester height measure and measured to the nearest centimetre. Readings of ≥ 0.5 cm were rounded up and < 0.5 cm rounded down to the nearest cm. Weight was measured with shoes and outdoor clothing removed in kilograms to the nearest 100g. Blood pressure, heart rate and pulse oximetry were measured using a Dinamap ProCare monitor (GE Healthcare, US) and portable oxygen saturation monitor.

2.4.1.2 Pilot

Height was recorded (to the nearest mm) on a single occasion following enrolment. Weight in kilograms was recorded to the nearest 100g and measured at each study visit. Body temperature was recorded with a single use thermometer (Tempa DOT™, 3M, UK) to the nearest 0.1 degrees C. Change in temperature post dosing was assessed according to the following approaches: maximum temperature post dosing, and objectively measured temperature 8 hours post start of dosing. Clinical examination at study enrolment and on dosing day was performed by a study doctor. At all other times, chest auscultation could be performed by either the study doctor or nurse.

2.4.2 Clinical laboratory blood tests and genotype

2.4.2.1 Run-In

Blood samples were taken at each visit for the clinical and research assays as summarised in Appendix 9.1. Outcomes discussed in this thesis (total white cell count, neutrophil count, C-reactive protein) were analysed by the hospital laboratories at the respective clinical sites according to usual hospital procedures. Analysis of the blood parameters from samples taken at each study visit was restricted here to routine clinical laboratory tests of inflammation, to allow comparison with the results following the single nebulised dose of gene therapy in the Pilot study (Chapter 4).

Patients with incomplete CF mutation information at study enrolment (i.e. either single or both mutations unknown) had an EDTA blood sample sent for CFTR mutation analysis (ARMS29 mutation screen).

2.4.2.2 Pilot

Blood for electrolytes, renal function, inflammatory markers, full blood count and coagulation screen were taken at study visits according to trial protocol. They were processed by the RBH clinical laboratories in accordance with usual routine hospital practice. A coagulation screen was only performed on subjects in the invasive cohort of the study, as a safety precaution prior to bronchial biopsy at bronchoscopy (these patients were all in Pilot Part A). Blood was also processed for cytokine analysis, autoantibodies and lymphocyte analysis of anti-*CFTR* T-cell response to investigate any host immunological response to the gene therapy product; however, the results are not included within this thesis.

2.4.3 Structural: Chest computed tomography (CT) scan

2.4.3.1 Run-In

A baseline CT chest was carried out at Run-In study visit 1 or 2 (or visit 3 in exceptional circumstances). Post menarchic female study participants were required to have a negative urine pregnancy test on the day of the CT scan. CT

scans performed for the Run-In study were volumetric high-resolution chest CT (inspiration, contiguous thin sections (1.25mm)) and interspaced high-resolution CT (expiration, 1mm sections at 10mm intervals) scanning from apices to bases. The maximal total dose of radiation for the CT scan was 1.5 milli Sieverts. CT scans were anonymised by coding and scored by two independent radiologists according to a scoring system developed by Professor D Hansell, Royal Brompton Hospital.

Scoring was carried out for the following variables: extent of bronchiectasis, severity of bronchiectasis, wall thickening, presence of small airway plugs, presence of large airway plugs and degree of air trapping. A score was given for each variable for each lobe (right upper, middle and lower lobes, left upper, lingual and lower lobes) by each radiologist. Where discrepancies of > 1 between scores existed, both radiologists reviewed the images and a consensus was agreed. Air trapping (on expiratory images only) was quantified as a percentage of the lobe involved to the nearest 5%. The final score entered into the CRF for each variable represented the mean for all lobes combined. CT scans were also assessed for the presence of ground glass or consolidation. The scoring system is summarised in Table 2.6.

Score	0	1	2	3	4	5
Extent of bronchiectasis	Nil	<25% lobe involved	25-50% lobe involved	51-75% lobe involved	76-100% lobe involved	
Severity of bronchiectasis	Absent	Trivial dilatation	> 1 but <2x diameter of adjacent blood vessel	2-3x diameter of blood vessel	>3x diameter of blood vessel	
Wall thickening	None	Trivial	Up to 0.5x diameter of adjacent blood vessel	>0.5x and up to diameter of blood vessel	>1 and up to 2x diameter of blood vessel	2x diameter of adjacent blood vessel
Small airway plugs	Absent	Mild	Extensive			
Large airway plugs	Absent	Mild	Extensive			
Air trapping	To nearest 5%					

Table 2.6. CT scoring grid used for analysis of Run-In chest CT scans. Scoring was performed by two radiologists, with a score given for each variable for each lobe (right upper, middle and lower lobes, left upper, lingual and lower lobes). Air trapping (on expiratory images only) was quantified as a percentage of the lobe involved to the nearest 5%. Final CT scores for each outcome (e.g. wall thickening) represented the mean for all lobes combined.

No other chest CT scans were performed for the Run-In study between visits 1-4; therefore there was only a single CT time point for data analysis for each patient. The second intended time point for CT in the Run-In study was either at a subsequent study 'exit' visit, or following recruitment to the multidose trial. As the latter is on-going at the time of writing this thesis, I have not included any data from paired scans.

2.4.3.2 Pilot

In part A of the Pilot study, patients underwent limited cut high resolution CT chest (without contrast) at intervals according to study protocol. Cuts were made every 30mm with seven images per CT scan. One pre-dose scan (<7 days pre-dose), and three post dose scans were performed (day 2, 14 and 28). The total number of chest cuts for the entire study was therefore 28 per patient. This was equivalent to an estimated maximum total radiation dose of 1.6 milli Sieverts. Female trial participants underwent urine pregnancy testing on date of each CT scan in accordance with usual clinical practice for radiological imaging.

CT scans were analysed by Professor Hansel (Consultant Radiologist, RBH) for signs of toxicity following nebulised gene therapy. As they were included as a safety measure (rather than efficacy), scans were not reported according to the scoring system described for the Run-In study.

2.4.4 Patient reported: Quality of life questionnaire

2.4.4.1 Run-In study

Quality of life was measured with the disease specific CFQ-R questionnaire (106;107), using the CFQ-UK (Revised) version (2005, Bryon & Stramik CFQ-UK English language version 1). This was completed at the start of each study visit. The CFQ-R measures functioning during the past two weeks and contains generic and CF specific responses. Children <14 years completed the child questionnaire. Adolescents and adults completed the adult questionnaire. Where transition across the 14th birthday occurred during the study, patients transferred to the adult questionnaire. All CFQ-R responses were converted to scores for the questionnaire domains by Dr Su Madge, Royal Brompton Hospital. According to questionnaire responses, a numerical score (out of 100) was given for each of the domains with higher scores implying greater quality of life. There are 12 domains in the adult and 8 in the child questionnaire, these are listed in Table 2.7.

CFQ-R Domains	
Adult and Child questionnaire	Adult questionnaire only
Respiratory	Role
Physical functioning	Vitality
Eating disturbances	Health perceptions
Treatment burden	Weight
Emotions	
Social limitations	
Body image	
Digestion	

Table 2.7. Adult and paediatric domains for the CF Quality of Life questionnaire (CFQ-R)

Scores from adult and child questionnaires were combined for analysis (108). Data analysis has concentrated on the Respiratory Symptoms domain. The questions from the CFQ-R which contribute to the Respiratory Symptoms domain score are displayed in Table 2.8. A minimal clinical difference for the Respiratory Symptoms domain of the CFQ-R in CF has been reported as 4 points (108).

Child version of CFQ-R questionnaire	Adult version of CFQ-R questionnaire
“Tell me how often in the past two weeks”: • You coughed during the day* • You woke up during the night because you were coughing* • You had to cough up mucus* • You had trouble breathing*	Indicate how you have been feeling during the past two weeks: • Have you been congested?* • Have you been coughing during the day?* • Have you had to cough up mucus?* How often during the past two weeks: • Have you been wheezing?* • Have you had trouble breathing?* • Have you woken up during the night because you were coughing?*

Table 2.8. Questions in CFQ-R contributing to the Respiratory Symptom domain score. Questions are shown according to questionnaire version (child or adult). Potential responses are on a four point Likert scale:

*Always/Often/Sometimes/Never. **A great deal/Somewhat/A little/Not at all.

2.5 Functional efficacy outcome measures: Potential difference measurements in the upper airway

This section describes the methodology for measurement of nasal PD in the Pilot study. Basic equipment required for nasal PD testing includes electrodes, a high impedance voltmeter, a perfusion set up, and data capture. The details of hardware vary according to manufacturer and test variability is also introduced with experimental technique (e.g. anatomical site of measurement) and experience of the operator. In both the US and in Europe, attempts have been made over recent years to standardise conditions between centres (94;109). However, although the technique is not difficult, it is one for which familiarity is essential and results can be affected by methodology such as times for solution

perfusion etc. Ahrens et al reported considerable variability in nasal PD results between centres (84) and therefore attention to detail in methodology and data analysis is essential.

The equipment set up at our centre (RBH) has been developed in our centre over the past 15 years. Several changes were required prior to trial onset, in order that it complied with all aspects of Good Clinical Practice. A Standard Operating Procedure was written by the author and followed by the researchers undertaking testing throughout the Pilot study (GD/JD), a detailed description of methods within this is given in section 2.5.2 below. A novel approach to nasal PD is also described.

2.5.1 Measurement of nasal PD: overview

Our methodology measures transepithelial nasal PD using a set-up involving an exploring electrode sited within an adapted Foley catheter and a reference electrode sited on the subject's forearm. The integration of a voltmeter into this 'circuit' permits measurement of PD. A photograph of the nasal PD set up is shown in Figure 2.9.

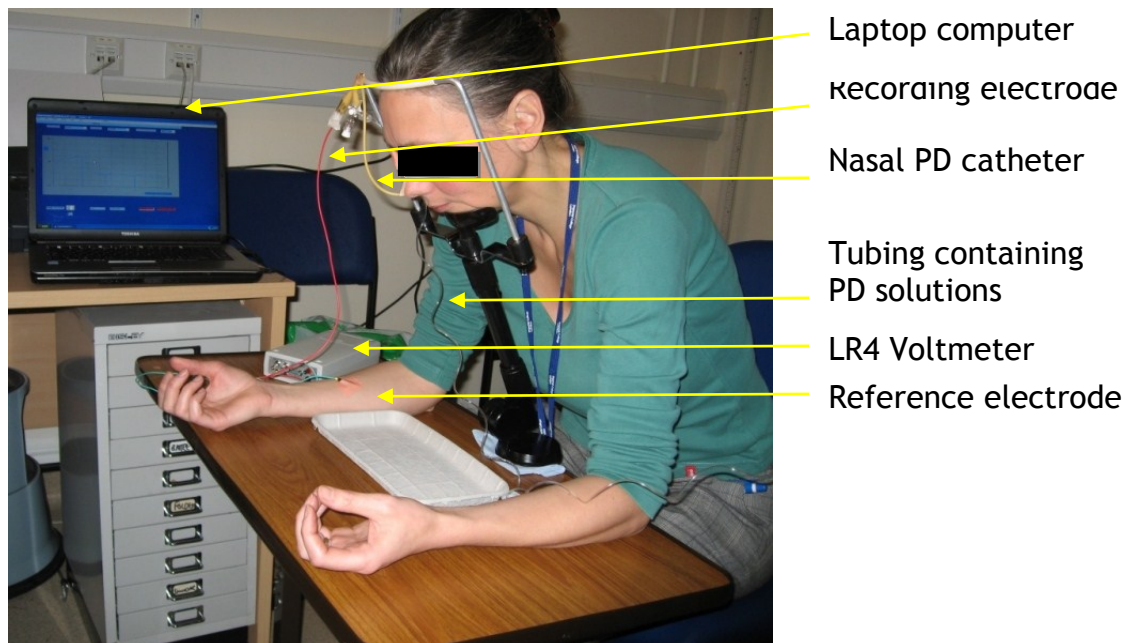


Figure 2.9. Nasal PD equipment set up. Photograph of nasal PD set up including nasal PD catheter with recording electrode, skin reference electrode on forearm connected to voltmeter and computer. Head position is maintained tilted slightly forward in head rest as shown.

2.5.1.1 Nasal PD measurements *in vivo*

The transepithelial PDs generated are detected by a voltmeter and displayed in real time on a laptop computer.

The nasal PD perfusion protocol includes the following solutions, the composition of which is to be found in Appendix 9.4:

1. Ringers solution (physiological; designed to reflect basal [non-stimulated, resting] PD)
2. Ringers with Amiloride (RA)
3. Zero Chloride solution with Amiloride (ZCA)
4. Zero Chloride solution with Amiloride and Isoprenaline (ZCAI)

Perfusion steps 1 and 2 are considered to measure sodium ion (Na^+) transport; steps 3 and 4 chloride (Cl^-) ion transport. Step 4 assesses *CFTR* specific (cAMP-activated) chloride transport. In CF, the basal nasal PD is higher (more negative), there is a profound depolarisation (becomes less negative) in the presence of amiloride (ENaC blocking agent), and absent or minimal evidence of Cl^- secretion in response to zero-chloride containing solutions and isoprenaline in contrast to the hyperpolarisation these solutions induce in non-CF individuals. Example representative traces from a non-CF and CF PD test are shown in Figure 2.10.

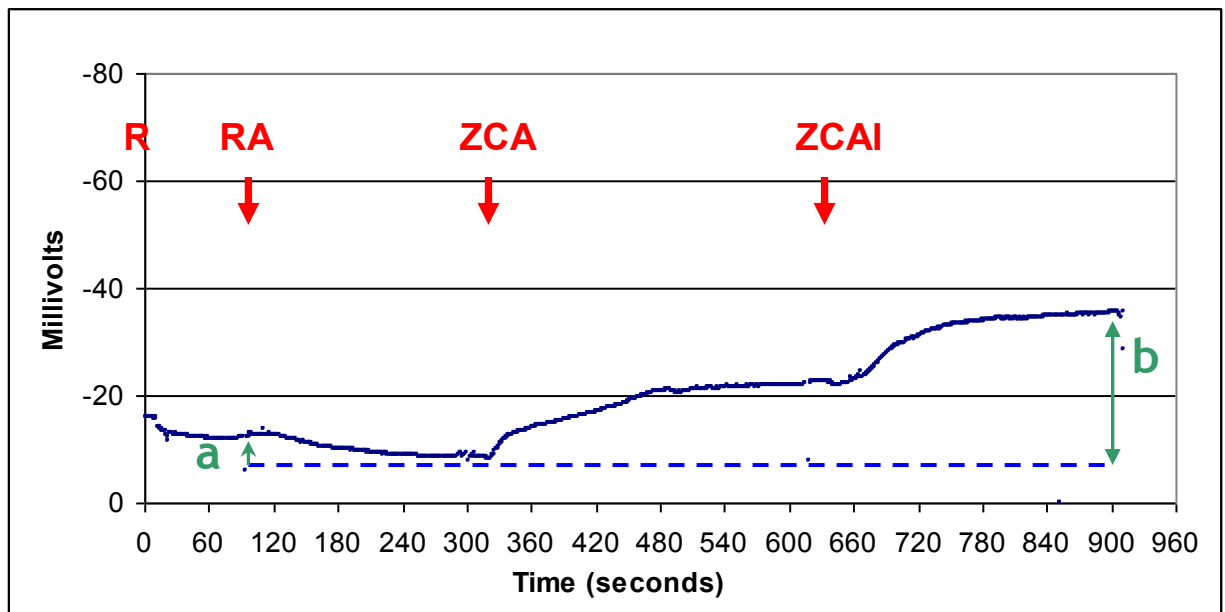


Figure 2.10A

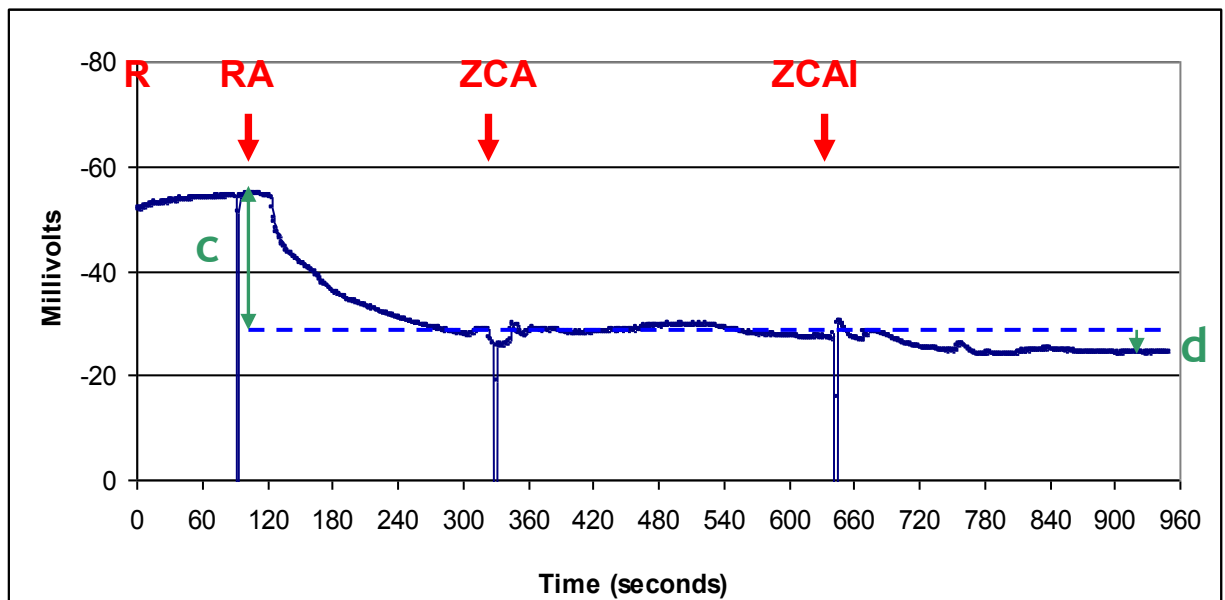


Figure 2.10B

Figure 2.10 Nasal PD traces in normal and CF subject.

Example nasal PD measurement trace from non-CF subject (2.10A) and CF patient (2.10B). Red arrows denote start of change in perfusion solution. R = Ringers solution, RA = Ringers amiloride, ZCA = zero chloride amiloride, ZCAI = zero chloride amiloride with Isoprenaline solution. Blue dashed line represents PD post RA perfusion, to compare with baseline PD during Ringers perfusion and post ZCA / ZCAI perfusion. Basal PD in CF is greater (more negative), as is change in PD post RA perfusion ($c > a$). PD increases in response to ZCA/ZCAI solutions in normal subjects (b) but remains flat or depolarises in CF (d). Vertical depolarisation spikes on CF trace are artefact and represent time points of solution change.

In addition to specific equipment and solutions, nasal PD requires technical expertise. Training on non-CF volunteer subjects was undertaken by GD prior to supervised (by JD, approximately 10 years of nasal PD experience) measurements during the Pilot study. Once competence had been achieved, all nasal PDs were performed by either JD or GD.

2.5.2 Patient testing and detailed description of set-up

All patients in Pilot part A and all patients in the nasal cohort of Pilot part B had nasal PD measurements performed. Intervals were determined by study protocol. This was open-label, designed to be iterative and underwent several amendments during the trial. In some cases, additional (off-protocol) tests were carried out to further quantify the duration of observed responses. Three nasal PD measurements were planned pre-dose of gene therapy because of the known variability within measurements between patients on different days. Post dose nasal PDs were performed until day 28, or longer in some cases if any PD changes persisted.

2.5.2.1 Preparation of Foley catheters

For the first time, there is now a commercially available double lumen catheter designed specifically for these measurements (<http://www.marquat.com/>). This was not available at the start of the trial and so our in house methodology of adapting a Foley catheter (Rusch, Teleflex Medical) was employed. Foley catheters have two channels but in their usual use as urinary catheters the purpose of one channel is to be able to inflate the balloon to keep the catheter in a patient's bladder. The catheter for nasal PD measurement requires two lumens throughout its length and the balloon is not required. Catheters were therefore adapted by the author by removing the balloon section and replacing the catheter tip, before creating a new side hole which allows an electrical connection between the surface epithelium and the 'exploring' lumen, which is filled with electrode cream (diluted 1:1 with Ringers solution) prior to use. The hole extends through the entire catheter, allowing the second lumen to perfuse solutions onto the nasal epithelium. Markings to a distance of 7cm (1cm intervals) from the side holes along the length of the catheter were made to guide appropriate placement during the repeated measurements made during the Pilot study. An 'adapted' Foley catheter is shown in Figure 2.11.



Figure 2.11. Photograph of adapted Foley catheter. There are two patent lumens throughout the length of the catheter. The first ('exploring') lumen is filled with electrode cream. An electrode is secured into place within this cream-filled lumen with tape which, for the purposes of illustration, is not shown. The second lumen is for solution perfusion, with a 3-way tap attached as shown.

For the first measurement in each subject, the PD catheter was inserted along the floor of the patient's nasal cavity and secured in place so that the exploring side hole sits at the site of highest (most negative) PD. Subsequent nasal PDs within the same subject were performed at the same site, unless there was a reason why this was not possible.

2.5.2.2 Solutions

To comply with Good Clinical Practice (GCP) and hospital pharmacy requirements, solutions for PD measurement were produced off site by an approved pharmacy manufacturing unit (St Georges Hospital, London). Prior to the Pilot trial, solutions had been made in-house and were based on a HEPES buffering solution.

Requirement for superior grade of reagents meant that the change was made to comply with the US TDN solution characteristics, based on Ringers solutions (110).

Stock solutions (in 100ml volumes, ethyl vinyl acetate (EVA) bags) were produced from the individual chemical reagents (Spectrum, US) and couriered to our centre in batches following production. Solutions were stored refrigerated prior to day of use, in a temperature mapped fridge. The stock solutions for nasal PD were Ringers, Ringers Amiloride and Zero Chloride Amiloride solution.

The effect of storage in plastic EVA bags was assessed by Quality Control testing carried out by the pharmacy production unit at St Georges Hospital. They

confirmed stability of storage of each stock solution type with refrigeration over a three month period (previously untested in EVA storage).

Solutions were removed from the refrigerator on the day of requirement and were allowed to reach room temperature before use. They were not warmed prior to perfusion. Each bag of stock solution could be kept for a period of 24 hours from opening prior to expiry (recommended from a microbiological rather than active substance perspective).

The fourth perfusion solution (containing isoprenaline) was prepared immediately prior to PD testing by adding isoprenaline (as isoprenaline sulphate) to a zero chloride amiloride stock solution. Solutions containing isoprenaline had an expiry time of two hours after preparation.

Documentation of solutions including batch number and expiry dates were completed on pharmacy-prepared worksheets for each nasal PD. A list of batch number and expiry date of each solution used was also placed in the patient's hospital medical records at conclusion of the trial, to comply with hospital pharmacy requirements for traceability.

2.5.2.3 Equipment

The high impedance, low resistance voltmeter used for making the PD measurements was model LR4 (Logan Research, UK). This was connected to a laptop computer (Toshiba) via a USB connection, with appropriate internal electrical isolation. The laptop power supply plug used was a Medline EXM80 80w AC/DC Wall plug adaptor (Powerbox, Craftec, UK). This was required in order to comply with safety regulations regarding appropriate electrical isolation for near-patient testing, with class II double isolation and approval for use with electrical medical equipment (standard IEC60601-1). Electrical safety testing involving the combined components of LR4 voltmeter, laptop and power supply was carried out by the BioMedical Engineering department at RBH. The electrodes (two per nasal PD) were silver/silver chloride (Ag/AgCl) 7mm discs (commonly used in electrophysiology studies including electroencephalograms (EEGs))(Biosense Medical, UK).

2.5.2.4 Offset and correction

The offset of a pair of electrodes used for measurement was checked prior to use in a nasal PD test. Offset describes the potential difference generated by creating a 'circuit' with the electrodes and conducting cream alone and is an indication of the suitability of electrodes for testing. To measure offset, electrodes were placed into a 20ml universal container containing 10ml electrode cream and 10ml Ringers (diluted 1:1) and voltage recorded by the LR4 as displayed on the laptop screen.

Because of the small mV values being recorded, tight acceptability criteria were used for electrode offset. Acceptability criteria were met if the (uncorrected) offset was $<5\text{mV}$. This offset was then corrected to zero using the Logan Research LR4 software ('offset correction' button within software). If an offset $\geq 5\text{mV}$ was obtained, the electrode pair was not used and a different electrode selected until an offset $<5\text{mV}$ was achieved.

2.5.2.5 Data recording

Patient ID, visit date, nostril used for testing and distance into nostril were recorded on a paper proforma. This was also used to record skin PD. Nasal PD data files were saved on the laptop computer and subsequently backed up to USB and network drive. No identifiable patient details were saved, files were saved according to by study ID, visit name and nostril side (e.g. 030D28R for patient 1SD030 day 28 visit, right nostril).

2.5.2.6 The measurement tubing 'circuit'

The tubing 'circuit' essential for perfusion of solutions onto the nasal epithelium was created using a series of plastic tubing connectors connected with 3-way taps. Each perfusion solution in a 50ml syringe was mounted in an infusion pump (Alaris[®] Syringe Pump) and connected to the tubing circuit, to allow the sequential perfusion steps to take place.

2.5.2.7 Preparation of nasal PD catheter

The adapted Foley catheters (prepared as described above), were single use and filled with the diluted electrode cream at the time of testing. The widest lumen of

the catheter was filled with diluted electrode cream (1:1 ratio with Ringers solution) using a 2ml syringe, taking care not to create air bubbles during filling. Air bubbles result in loss of electrical conductivity and disruption to the measurement 'circuit'.

The Ag/AgCl disk electrode (the 'exploring' electrode) was inserted into the same lumen of the cream-filled catheter at the large port end, and taped with Slek to secure in position and prevent escape of electrode cream. The electrode plug was connected to the positive (red) port of the LR4 voltmeter as described previously.

2.5.2.8 Forearm (reference) electrode

It has been shown that the subcutaneous space in the forearm is isoelectric with the submucosal space of the airway (111). Access to this can be achieved by removing the epithelial skin layer; at our centre this is achieved by skin abrasion. We do not use the agar filled needle inserted subcutaneously into the forearm for the reference electrode utilised in the US TDN nasal PD protocol (110).

The skin was cleaned using an alcohol wipe. A skin abrasion was made by carefully using a sterile dental drill burr. This was applied to the forearm surface until a sheen on the skin was seen. A piece of tape (slek) was placed over the reference electrode and a hole made through the tape into electrode with the tip of the dental burr. This was placed exactly over abraded skin site and secured using tape. A 2ml syringe was filled with diluted electrode cream, and this was 'injected' into a hole in the tape (created using a needle sheathed with a sterile pipette tip attached to the syringe) to completely fill the dead space between electrode disk dome and skin. This created good electrical contact between the abraded skin and electrode itself.

The skin electrode was connected to the negative (black) port of the LR4 voltmeter as described above.

2.5.2.9 Skin PD

Following identification of a satisfactory pair of electrodes by offset testing and catheter/tubing set up prepared, and siting of forearm electrode, skin PD was

measured. This was undertaken by placing the prepared nasal PD catheter (with recording electrode attached) on the index finger tip. This enabled confirmation that the integrity of the electrical circuit and skin contact was good. Measurements *in vivo* are typically between -30 and -60 mV (with values in CF > non-CF); and if skin PD was <30mV the circuit was reviewed and forearm skin preparation process repeated.

Skin PD was confirmed prior to commencing nasal measurements for each subject on each test occasion.

2.5.2.10 Nasal measurement

The nasal PD test was performed with the subject seated at a table (adjustable height) and we incorporated a chin rest (Double screw clamp chin rest, 6011R, Richmond Products, NM, US) into the set up to aid head stability and comfort during testing. The head rest was adjusted to ensure that the head tilted slightly forward so that solutions dripped out of the patient's nostril rather than trickled down their nasopharynx.

To insert the nasal PD catheter, the subject was asked to lift their head off the chin rest so that the researcher had an optimal view of nasal passages. The PD catheter was inserted slowly along nasal floor whilst watching mV readings in real time on computer screen. The final site of catheter placement was that at which the PD was highest (most negative). The nostril and distance inserted was recorded (in centimetres). The PD catheter was taped in place using tape strips (Steri-Strip™, 3M, Bracknell UK) wrapped around the catheter and taped to the nose. Additional catheter site security was provided by taping the catheter to the head rest. The same nostril and distance into nose was used for all subsequent measurements involving that subject, unless readings could not be achieved at subsequent tests. On the rare occasion this occurred (e.g. unable to tolerate catheter in particular nostril due to pain) it was documented.

2.5.2.11 Solution perfusion

Prior to commencing the sequence of solution perfusion the subject's head was positioned on the chin rest. A drip tray was placed on the table to retain used

perfusate. Solutions were perfused according to the following timescale at a flow rate of 4ml/minute (240ml/hour), in accordance with previously described optimal flow rates for testing:

- Ringers until stable (<1mV change in PD) for 30-60 seconds
- Ringers Amiloride for at least 3 minutes but until stable (defined as above)
- Zero Chloride Amiloride solution for 5 minutes
- Zero Chloride Amiloride with Isoprenaline for 5 minutes

To mark the time point at which a solution change occurred, the calibration switch on the voltmeter was used as a marker by creating an artefact reading. This was easily identifiable on the computer screen as a large spike on the PD trace out of keeping with true transepithelial measurements.

Any other relevant details, e.g. patient coughing or moving during procedure were recorded on the paper proforma completed for each PD measurement.

Millivoltages detected by the LR4 were displayed graphically in real time whilst testing. Each PD test continued until the sequence of perfusions was complete, unless patient factors or technical factors (e.g. loss of electrical contact) precluded test completion.

2.5.2.12 Infection control

Royal Brompton Hospital policies on infection control were followed for patient isolation and testing equipment. A new Foley catheter was used for each test and disposable plastic tubing used for solution perfusion. Hard surfaces including the chin rest used for head placement during testing were cleaned with disinfectant wipes (Clinell Universal Wipes, Steroplast Healthcare, UK) between patients. An electrode was designated the skin electrode for each patient for the duration of the trial, with no cross contamination of skin electrodes between patients. If an existing skin electrode did not comply with $\leq 5\text{mV}$ offset at subsequent test dates, a further new skin electrode was designated to that patient.

2.5.2.13 Generating data files

Files generated from the Logan Research LR4 software were prop.binary data files with extension (.dat). Text files were created by clicking on 'DataDump' in the file menu within the LR4 software, automatically creating a file with extension .txt. These files could be imported into Microsoft Excel, for subsequent data analysis.

2.5.3 *Methods: data analysis*

Formal nasal PD data analysis took place following completion of data collection in the Pilot study. Nasal PD traces were analysed independently by the author (GD) and JD prior to combining the analysis as described below.

2.5.3.1 Generating text export files for analysis

Raw data files (.txt) were imported into Microsoft Excel. Each row of data on the spreadsheet represented the mV recorded in 1 second intervals from start of data capture to end of recording.

2.5.3.2 Blinding of test identification

During PD analysis, GD and JD were blinded to patient identification and timing of test including whether the traces represented pre- or post dosing measurements. The author achieved this by attributing a random 4 digit code number to the PD test and removing all features relating to patient study code and visit date. Random numbers were generated using the Random Number generator in Microsoft Excel.

2.5.3.3 Analysis parameters

Export of files to spreadsheet in Excel allowed detection of the mV and corresponding time point (in seconds) following variables that contributed to analysis:

- Basal PD:

Basal PD was identified according to a period of stability on the PD trace during perfusion with Ringers solution. The value taken for basal PD represented the average (mean) of 20-30 seconds' mV values.

- Amiloride response:

The amiloride response was defined as the delta between the start and end of Ringers Amiloride perfusion. The start value was taken as the basal PD unless the PD following marker artefact (created using calibration switch as described above) differed substantially from this. The mV at the end of RA perfusion was taken where the mV had plateaued out, again taking the average of around 30 data points. The delta for Ringers Amiloride response was the difference between the start and end of Ringers Amiloride perfusion.

- Zero chloride amiloride (ZCA) response

The zero chloride amiloride response was defined as the delta between the start and 5 minutes post start of ZCA perfusion. The starting mV was taken as the plateau of the amiloride response unless the PD immediately following marker artefact (to mark switch to ZCA solution) substantially differed from this. The mV value at the end of 5 minutes perfusion was the average of preceding 30 data points. The delta was the difference between the mV at the start and end of ZCA solution.

- Zero chloride amiloride isoprenaline (ZCAI) response

This was calculated using the same methodology as described for ZCA response, for the ZCAI perfusion stage. The sum of the change in mV (delta ZCA + delta ZCAI) was also calculated, representing total chloride transport.

2.5.3.4 Combining two nasal PD analyser results

Nasal PDs from patients receiving the preliminary 1ml nasal dose (n=3 patients) were analysed solely by GD as this was a small safety subgroup which preceded the

full trial. All nasal PD results from patients receiving the 2ml nasal dose were analysed by GD and JD. Following independent analysis, analyses were then combined to determine 'final agreed' PD values for each variable of interest.

To determine final agreed PD values, acceptability criteria were pre-defined for maximal allowed differences between the two analysers' values. The final agreed mV value for each parameter was determined as follows:

- The differences between values calculated by GD and JD for each analysis parameter were calculated. Values for each analyser were automatically accepted if the difference between GD and JD were within the following range (in millivolts (mV)):
 - Ringers: 2mV
 - Ringers amiloride delta: 2mV
 - Zero chloride amiloride delta: 1mV
 - Zero chloride amiloride isoprenaline delta: 1mV
- Tests where the difference between GD and JD was greater than these maximum values were discussed. Where, following review of plotted values, both JD and GD values were considered equally valid, the mean of both values was calculated.
- For tests where disagreement out with the above criteria existed, PD traces were reviewed by JD and GD and a consensus opinion was agreed about final values or requirement for further review.
- Tests that were deemed non-analysable by both analysers were excluded from the final values for data analysis.
- Where a disagreement out with the above criteria persisted following repeat analysis, PD traces were reviewed by JD and GD and a consensus was agreed about final values.

The final mV values for each of the PD analysis parameters for each nasal PD were calculated by obtaining the mean of GD and JD analysis values.

Zero chloride responses (deltas) were calculated for the ZCA stage and ZCAI stage separately, and also by summing the two deltas for a mV result to represent total chloride secretion.

2.5.3.5 Pre-dosing tests

The Pilot protocol included more than one pre-dose nasal PD per patient, with all patients scheduled to have at least two, and up to 3, measurements pre- nasal gene therapy. In order to benchmark post dosing measurements to those taken pre-dose, mean pre-dosing values were calculated. These were calculated using the final agreed values (calculated as per GD/JD as above) to obtain a mean value for all pre-dosing tests per patient. Average pre-dosing measurements were calculated for the following variables: basal, and deltas for amiloride, zero chloride amiloride and zero chloride amiloride isoprenaline responses to allow comparison with post dose results. Each individual pre-dose test result was also considered individually for the calculation of chloride secretion ‘response’ according to the methods described in section 2.5.3.6 below.

2.5.3.6 Defining efficacy response

There is no ‘gold standard’ method to define change in sodium or chloride ion transport in response to a novel therapeutic intervention such as gene therapy. For each section of the nasal PD tests, within-patient variability in pre-dose tests was investigated.

The optimal way of classifying evidence of chloride secretion on nasal PD results is unknown. It was the responses to zero chloride with amiloride (ZCA) solutions that were shown to change post CF gene therapy in the Lancet study (35). For Pilot results, intrinsic chloride secretion (ZCA delta), stimulated chloride secretion (ZCAI delta) and total chloride secretion (sum of ZCA and ZCAI deltas) were calculated for each PD test. Evidence of chloride transport on post dosing nasal PDs was then analysed by the author in accordance to the following three definitions:

Method 1: Evidence of chloride secretion on post dose test greater than any pre-dose test for that individual

For individual patient results, the method involved defining ‘response’ according to evidence of chloride secretion on a post dose test greater than any pre-dose test for that individual. This was investigated for intrinsic chloride secretion (ZCA delta), stimulated chloride secretion (ZCAI delta) and total chloride secretion (ZCA delta + ZCAI delta) for each test.

Method 2: Chloride secretion relative to mean pre-dosing results

Response was defined as a post-dose test with a sum representing chloride transport $\geq 5\text{mV}$ relative to mean pre-dosing test results for that individual. This was performed for intrinsic chloride secretion (ZCA delta), stimulated chloride secretion (ZCAI delta) and total chloride transport (ZCA delta + ZCAI delta). Total chloride transport $> 5\text{mC}$ relative to pre-dosing tests has been used previously in CF clinical trials (96;97).

Method 3: ‘Normal’ chloride secretion

Finally, chloride secretion was assessed according to whether $\geq -5\text{mV}$ of total chloride transport was achieved (irrespective of the magnitude of pre-dose results). This cut off was used for total chloride transport as it is considered to represent normal (non-CF) levels of chloride secretion (112). In addition to performing this for total chloride transport (ZCA delta + ZCAI delta), results were analysed according to having $\geq -5\text{mV}$ in intrinsic chloride secretion (ZCA delta), and similarly for stimulated chloride secretion (ZCAI delta).

2.5.3.7 Relationship with other Pilot outcome measures

The relationship between change in nasal PD and the presence of vector specific *CFTR* mRNA on nasal brushings was assessed, with methods as described in section 2.7. Relationships with genotype and other patient characteristics were also investigated.

2.5.4 Statistics

Intra-class correlations (ICC) were calculated to assess inter-reporter agreement between nasal PD values for each section of the test (e.g. basal PD) as calculated by GD and JD. This was undertaken in several phases:

- Initial agreement between GD and JD for all 2ml nasal PD analysed values
- Final agreement between acceptable tests following consensus between GD and JD

ICCs were calculated using SPSS Statistics version 21.0 (IBM), using a two way mixed effects model for absolute agreement. Final PD values for each section of PD (mean of accepted GD and JD result) were combined to analyse group differences between pre- and post- measurements. Comparison between groups was performed using paired t-tests or Wilcoxon signed rank tests with significance at $p < 0.05$.

Time points for post dosing tests were determined by trial protocol. In addition, it became apparent early in the trial that if a change was seen at last time point defined in the protocol, it would be useful for further off-protocol measurements to be made until the end of trial (day 28 post dosing, or post day 28 in the presence of any persisting evidence of change).

Mean-pre and a single time point post dosing measurements (e.g. day 14) for basal, and deltas for ringers amiloride, zero chloride amiloride and zero chloride amiloride isoprenaline were compared using paired Students t-tests or paired non-parametric tests.

2.5.5 Patient inclusion in nasal dosing cohort

At the start of the Pilot study, we had less strict inclusion criteria for the nasal dosing cohort in terms of evidence of chloride secretion on pre-dosing nasal PDs. A minority of patients recruited early in the trial process had $\geq 5\text{mV}$ chloride secretion on pre-dosing tests but were not excluded from nasal dosing, yet we adopted a stricter policy for inclusion in the latter stages of the trial once the safety

parameters were becoming clearer in an effort to further define the nature of PD results.

2.5.6 Novel approach to nasal PD testing

A novel approach to PD testing was investigated which was hypothesized to improve the quantitative detection of change in PD in response to CF gene therapy.

2.5.6.1 Novel voltmeter and clamp

Standard nasal PD measurements take advantage of the change in electrical potential difference as a result of a chemical load. As described for the nasal PD perfusion protocol in the Pilot study, a change in PD is measurable as a result of exposing the nasal epithelium to the sodium channel blocker amiloride. This can be considered a chemical challenge or ‘load’ for the airway cells at the site of measurement. An alternative approach would be to subject the airway to an electrical challenge or ‘load’. The potential advantages would potentially include being able to apply and then remove the ‘load’ instantaneously to evaluate response within seconds and negate the need for perfusion stages in the PD protocol.

The hardware for a unit to evaluate this hypothesis was developed by Logan Research Ltd (Rochester, UK). It incorporates a voltmeter set-up similar to that used for standard nasal PD testing, but has an additional capability of being able to apply a high resistance ‘load’, of variable magnitude, at intervals during testing to the PD generated across the respiratory epithelium. This artificial manipulation of the PD essentially applies a resistance in parallel to the (parallel) electrical circuit created by the PD across the epithelial cells themselves. No additional voltage or ‘electricity’ is being given to the patient, rather the clamp load(s) are applied and recovery to baseline PD is assessed.

It was hypothesised that the magnitude of PD change in response to an electrical load and rate of return to baseline PD may be different depending on the proportion of cells expressing *CFTR*. It was considered that following the application of a high resistance load to the ‘circuit’, the proportion of cells with

functioning chloride channels would determine the rate of recovery of PD to pre-clamp values.

This method has never been applied *in vivo* to respiratory epithelial PD measurements. As a novel standalone piece of equipment, the hardware voltmeter-based unit ('LR5') was tested by the BME department at the RBH to ensure it complied with electrical safety testing prior to testing on human subjects *in vivo*. In addition, as it has no CE mark, standard procedures were followed to permit this unit to be evaluated within the Royal Brompton and Harefield NHS Foundation Trust.

2.5.6.2 Experimental set up

The LR5 was attached to a standard laptop computer (Toshiba, Windows XP operating system) via USB connection. A double insulated power pack (Craftec) replaced the proprietary laptop mains AC adaptor to comply with electrical safety testing appropriate for near-patient use. Frequency of signal filtration was manually selected on the front panel of the LR5. This included the following: Notch (50Hz) and WB wideband 1KHz low pass, Wide band (no notch filtering) 1 KHz low pass, Medium filter 10Hz low pass (F1 setting on LR5), Low filter 1Hz low pass (F2 setting on LR5). PD (in millivolts) was recorded every 0.03 seconds (33Hz). The clamp was applied during basal PD measurement in the absence of solution perfusion, with a load between '1' and '9'. Setting '1' represented no load; (maximal resistive component) 9 (maximal load; lowest resistive component). Clamp duration was manually adjustable between 1 - 10 seconds. The point at which the clamp is applied was seen graphically on the laptop screen in real time, as a pale blue line indented at 1 second intervals. This line has been set arbitrarily by the manufacturer at -69mV, and decreases to -59mV when the clamp is applied.

The electrode set up for the LR5 was similar to the usual LR4 voltmeter used for our existing nasal PD set up, however the LR5 uses a 3-electrode rather than 2-electrode set up. Two silver/ silver chloride electrodes were used to make electrical measurements across the respiratory epithelium. An electrode overlying abraded skin on the forearm was connected to the negative port of the LR5 as per conventional testing with the LR4 voltmeter. The nasal electrode (either in direct contact with airway surface or within an electrode cream filled Foley catheter

depending on experimental set up) was connected to the positive port of the LR5. In addition, a press-stud electrode attached to the anterior chest wall with a standard ECG electrode gel pad (Skintact® CT-601, Leonhard Lang Ltd) was used to provide an electrical ground reference. At the end of each test, files were saved as a data file extension (.dat) file. A corresponding ascii text file was generated by clicking 'datadump' within the LR5 software programme. These were imported into a Microsoft Excel spreadsheet for analysis.

2.5.6.3 Clinical investigation of LR5 device

To investigate the ability of the LR5 to manipulate the transepithelial PD, the magnitude of load, clamp time, signal filtration and type and set-up of recording electrode were assessed.

Measurements across the nasal epithelium and skin were initially made in healthy volunteers (n=2) and in CF subjects (n = 4) taking part in the Pilot gene therapy trial. The CF patients were enrolled into the nasal cohort of the gene therapy study and timing of LR5 measurements coincided with standard nasal PD testing as determined by trial protocol.

Comparison between non-CF and CF nasal PD responses to electrical load were made by analysing rate and magnitude of change in PD and recovery to baseline PD between subjects (expressed as % pre-clamp values). As stated, CF patients were evaluated with the LR5 post standard nasal PD testing. Therefore their respiratory epithelium had been exposed to the sequential perfusion steps with Ringers, Ringers amiloride and Zero chloride amiloride solutions immediately prior to LR5 evaluation. Due to the preliminary nature of the investigation, the author accepted this as a limitation of testing. No pre-perfusion tests with the LR5 were carried out on CF patients in order to maintain constant experimental conditions for the standard nasal PD as an outcome measure in the Pilot study. However, testing with the LR5 in non-CF subjects (n=2) both pre- and post standard nasal PD testing (with perfusion protocol) was undertaken to ensure this did not influence the responses seen.

2.6 Functional efficacy outcome measures: Potential difference measurements in the lower airway

Although any functional *CFTR* assessment would ideally be performed repeatedly, this is not feasible in the lower airway for both practical and ethical reasons. The nose is often used as a surrogate because of similarities in respiratory epithelium and accessibility to testing. However, there are differences with regards to drug delivery (which is much easier in the nose than in the airway, particularly when the latter is diseased) and local environment (e.g. presence of airway secretions) that may affect measurement. Direct PD measurement in the lower airway is thus relevant and important, despite the limitations of infrequent testing. Such measurements are technically more challenging than in the upper airway. Prior to the Pilot study, although lower airway measurements had been performed, these techniques would not have complied with current standards of Good Clinical Practice (GCP). Therefore initial work concentrated on methodology to develop an appropriate measurement set up (designed largely by the author), which was then utilised for *in vivo* measurements in the Pilot study.

2.6.1 Previous methodology and rationale for changing existing set up

Lower airway PDs have previously been undertaken within our department in adults (35) and children (113) and the method now described aimed both to address GCP compliance issues and to overcome some of the problems previously identified. The group originally used a double lumen PD catheter through the suction channel of an adult sized bronchoscope (35). One lumen was used for PD solution perfusion and the other was primed to the tip with Ringers and used as an electrical bridge to connect with the exploring electrode for lower airway PD measurements. This fluid interface required the use of glass Calomel half cell electrodes. The reference electrode in this set up was connected to an intravenous catheter. This set up was similar to the original methodology described by Knowles et al (74) for measuring upper (and lower) airway PD.

The bronchial PD set up described by Davies et al (113) included modification from a double to single lumen catheter to fit down the smaller channel of the paediatric

bronchoscopes. A single lumen polyethylene catheter was primed to the tip with a zero chloride solution. Basal PD measurements were recorded through this fluid column to the exploring electrode, prior to perfusion with a low chloride solution to assess chloride secretion.

Both of these techniques (35;113) provided technically acceptable readings however there were several relevant technical and regulatory issues. Use of Calomel half cell electrodes was problematic particularly due to microbubbles interrupting the fluid filled conduction circuit and the glass electrode interface required for PD measurement. Both set ups also required the presence of an intravenous line to connect with the reference electrode, again at a fluid-glass interface which needed to be bubble free. There were also considerable issues with meeting appropriate equipment sterilisation requirements. Therefore, for the current CF gene therapy Pilot study, circuit modification has been undertaken.

We chose to use the same silver/silver chloride (Ag/AgCl) 7mm disc electrodes as used for nasal PD measurement (Chapter 2, section 2.5). In addition, we were no longer able to source simple sterile double lumen or single lumen polyethylene tubing of clinical grade. Instead, we used a single lumen catheter sourced by using the tubing from the casing of a standard bronchoscopy cytology brush (the brush was discarded)(Olympus, UK).

No system for measuring lower airway PD with solid Ag/AgCl electrodes with the inclusion of basal PD measurements and a perfusion protocol has previously been described. Fajac et al (114) describe the development of a Ag/AgCl electrode technique for measuring PD in the human lower airway but their technique allows no perfusion so only basal measurements can be obtained. In previous CF gene therapy clinical trials, it is the response to zero or low chloride solutions perfusion that has been shown to change (rather than basal measurements) and therefore it was critical to develop an experimental set up and measurement protocol for assessing bronchial PD which could measure chloride secretion.

2.6.2 Methods and initial early phase results

2.6.2.1 Early modification to bronchial PD set up: electrodes

Incorporation of Ag/AgCl electrodes into the PD set up required consideration of the interface between the electrode and the remainder of the PD ‘circuit’. Whereas Calomel electrodes form a fluid interface by direct contact of their glass wall to a fluid column, Ag/AgCl electrodes are solid metal discs. Therefore, a system to ‘connect’ the Ag/AgCl electrodes with the fluid column measuring lower airway PD was required. The author chose to investigate this using agar as an interface. 4% agar Noble in 3 molar (M) potassium chloride (KCl) was chosen because it is often used as an electro- conductive bridge for *in vitro* and *in vivo* transepithelial PD measurements (110;115). In seeking to optimise this, the following conditions were explored:

2.6.2.1.1 Condition 1: electrode embedded in agar

- The electrode offset was checked on a pair of Ag/AgCl electrodes; accepted if <5 mV (Figure 2.12).

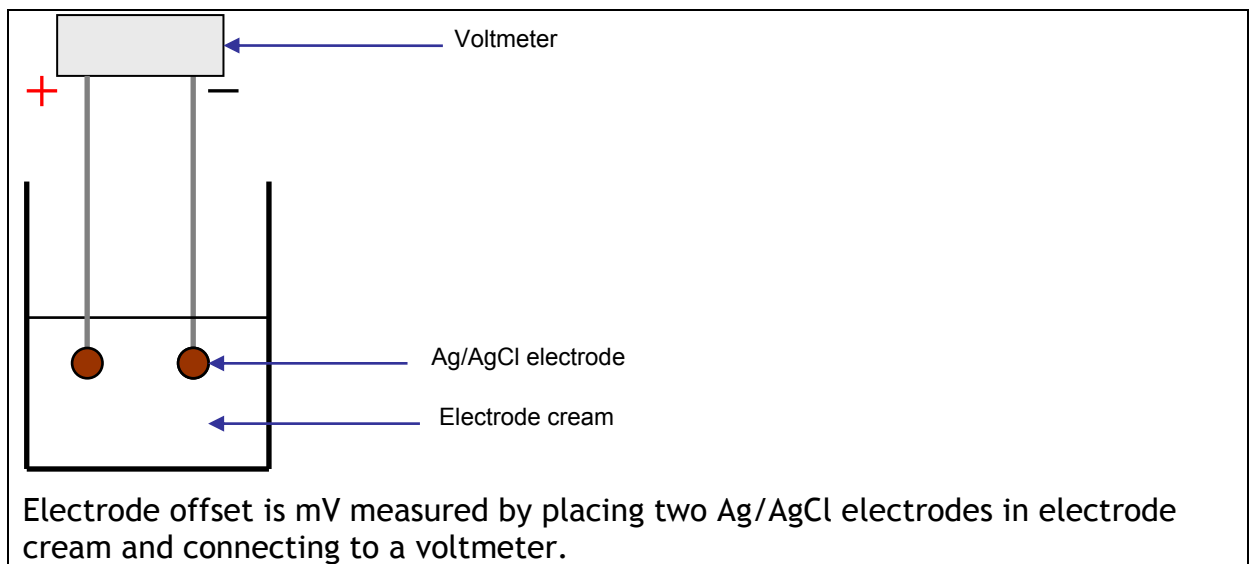


Figure 2.12. Testing electrode offset between two Ag/AgCl electrodes.

- One of the pair was embedded in 2ml 4% agar in 3M KCl in a 5ml syringe (with plunger removed); this was allowed to solidify at room temperature for four hours (n=4 tests).
- Circuit offsets were retested by placing the syringe tip containing solid agar into Signacreme (electrode cream)(Parker Laboratories, US). The free end of the embedded electrode was inserted into the positive port of a high impedance voltmeter. The second Ag/AgCl electrode was placed directly into the electrode cream, and into the negative port of voltmeter (Figure 2.13).

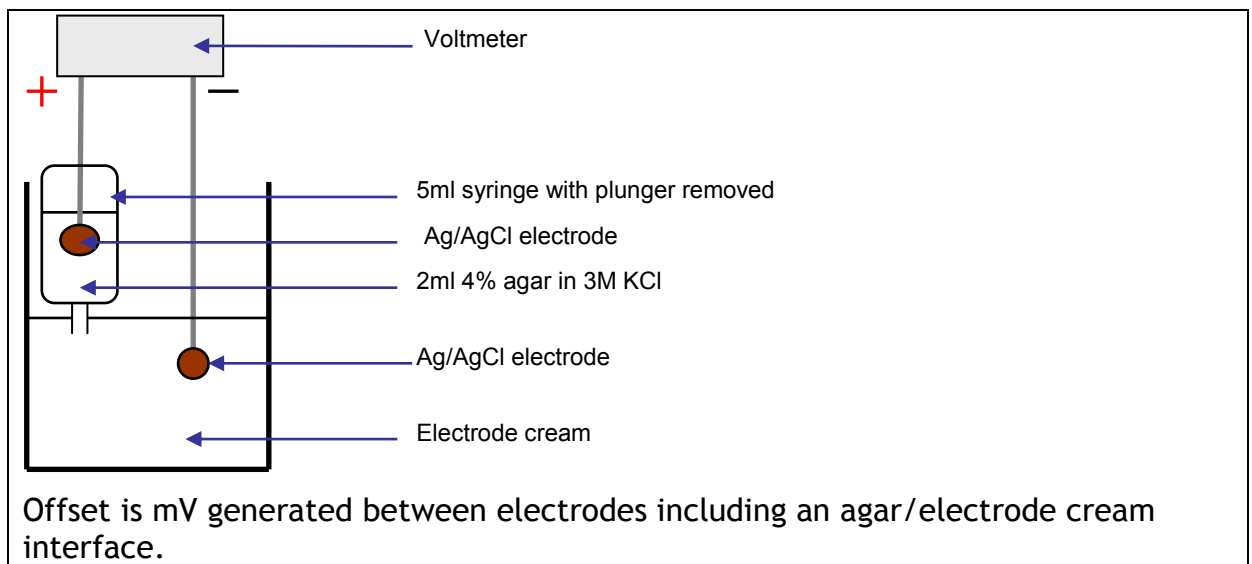


Figure 2.13. Development of electrode interface for bronchial PD circuit: testing offsets with electrode embedded in agar.

- Electrode offset was tested and defined as the difference (mV) between the two electrodes.

Outcome: Offsets were high (mean -48.0 mV, SD 8.0 mV). This indicated a significant conduction problem within the circuit, so whilst it would be technically possible to adjust for this, such an adjustment would lead to sensitivity problems in subsequent measurements.

2.6.2.1.2 Condition 2: agar and conducting cream interface

The author considered that the contact between the solidifying agar and the electrode was likely suboptimal and responsible for this problem. The author next tested a system she devised involving both agar and electrode cream (repeated for n=4 tests).

- 1ml volume of 4% 3M KCl agar was inserted (when liquid) into a 5ml syringe with the plunger removed.
- Once this had solidified, electrode cream (approximately 2ml) was placed on top of the agar
- The ‘exploring’ Ag/AgCl electrode was placed in to this electrode cream and secured in place.
- A separate Ag/AgCl electrode was placed directly into a separate pot of electrode cream, and into the negative port of voltmeter. The tip of the syringe containing agar (with exploring electrode) was also placed in to this pot of cream (Figure 2.14).

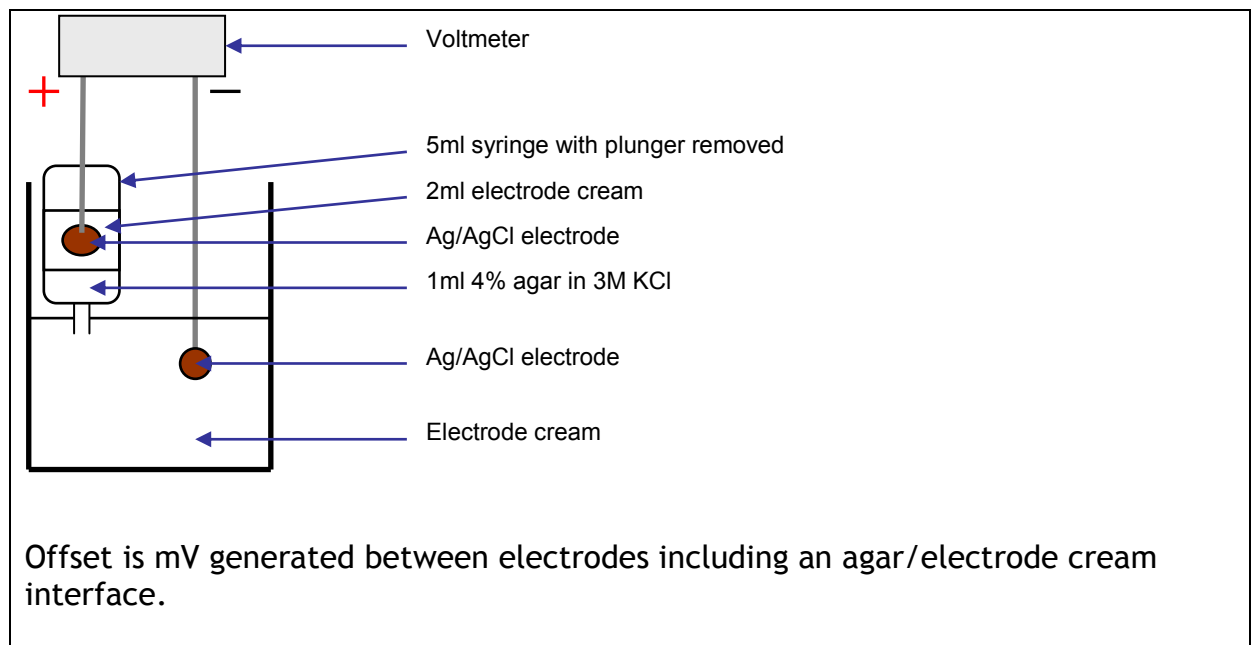


Figure 2.14. Development of electrode interface for bronchial PD circuit: testing offsets with electrode placed in layer of electrode cream on top of agar.

- Electrode offset was tested and defined as the difference (mV) between the two electrodes.

Outcome: Offsets were not affected by the presence of agar column (mean offset = -2.4mV, SD 0.7 mV) and were identical to those achieved when both electrodes were placed in electrode cream alone.

This electrode set up was subsequently used for all bronchial PD measurements in the Pilot study. Connection to the remainder of the PD circuit and overall offsets for the entire circuit are described in the following sections.

2.6.2.2 Initial bronchial PD set up: tubing and connections

The voltmeter (LR4), computer and software used for lower airway measurements were the same as those used for upper airway measurements in the Pilot study (section 2.5).

Solutions for bronchial PD are listed in Appendix 9.5 and the perfusion protocol is described below. Stock Ringers and Zero chloride solutions were stored refrigerated until use. For bronchial PDs undertaken with this ‘initial’ set up, solutions were taken out the refrigerator on the morning of use.

The tubing connections, cytology brush sheath and other components forming the bronchial PD circuit are shown in Figure 2.15. Preparation of the circuit involved:

- Removal of bronchoscopy cytology brush from casing; brush discarded and casing retained.
- Joining T-piece connectors with the use of 3-way taps to permit inclusion of a perfusion step with zero chloride isoprenaline solution as well as basal measurements.
- Priming a T-piece connector adjacent to a 3-way tap with zero chloride solution (without isoprenaline).
- Connecting the agar containing syringe to this primed T-piece connector.

- Adding 2ml Signacreme electrode cream to agar as described above
- Placing Ag/AgCl exploring electrode disc in Signacreme.
- Priming the remaining measurement circuit to the tip of the cytology brush sheath with Ringers solution.
- Performing basal measurements and first of three distal measurements before changing perfusion solution to zero chloride with isoprenaline.

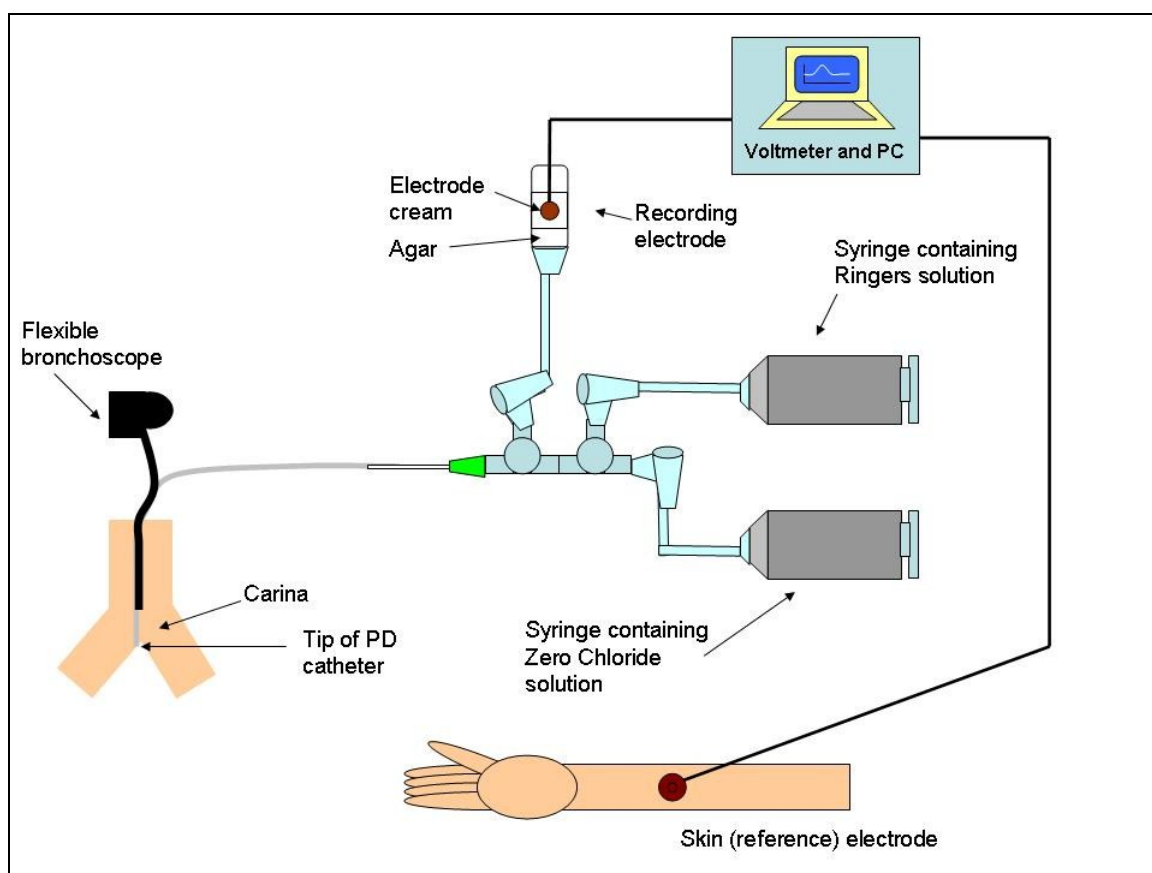


Figure 2.15. Schematic representation of initial bronchial PD set up. PD catheter (cytology brush sheath) is connected to two three way taps via an Abbocath catheter. T-piece connectors are used to connect the Ringers and Zero Chloride solution syringes, and the syringe containing Agar and electrode cream. Transepithelial measurements are made across the bronchial epithelium with the recording electrode embedded in electrode cream. An Agar-fluid interface conducts through the Ringers fluid column within the PD catheter to the apical surface of the respiratory epithelium down a flexible bronchoscope. The single channel PD catheter is also used to perfuse Zero Chloride (with isoprenaline) solution during continual PD measurement, with three-way taps facilitating change from Ringers to Zero Chloride solution. Reference electrode is placed over an area of abraded skin with contact optimised with electrode cream. Recording and reference electrodes are connected to a voltmeter and PD is recorded on a laptop computer.

2.6.3 Evaluation of the novel set up and use in Pilot study

Circuit integrity and ability to perform measurements in the nose of healthy volunteers (n=2 subjects, 3 perfusions undertaken) using the bronchial PD set up was confirmed prior to the first trial patient bronchoscopy with a representative example shown in Figure 2.16.

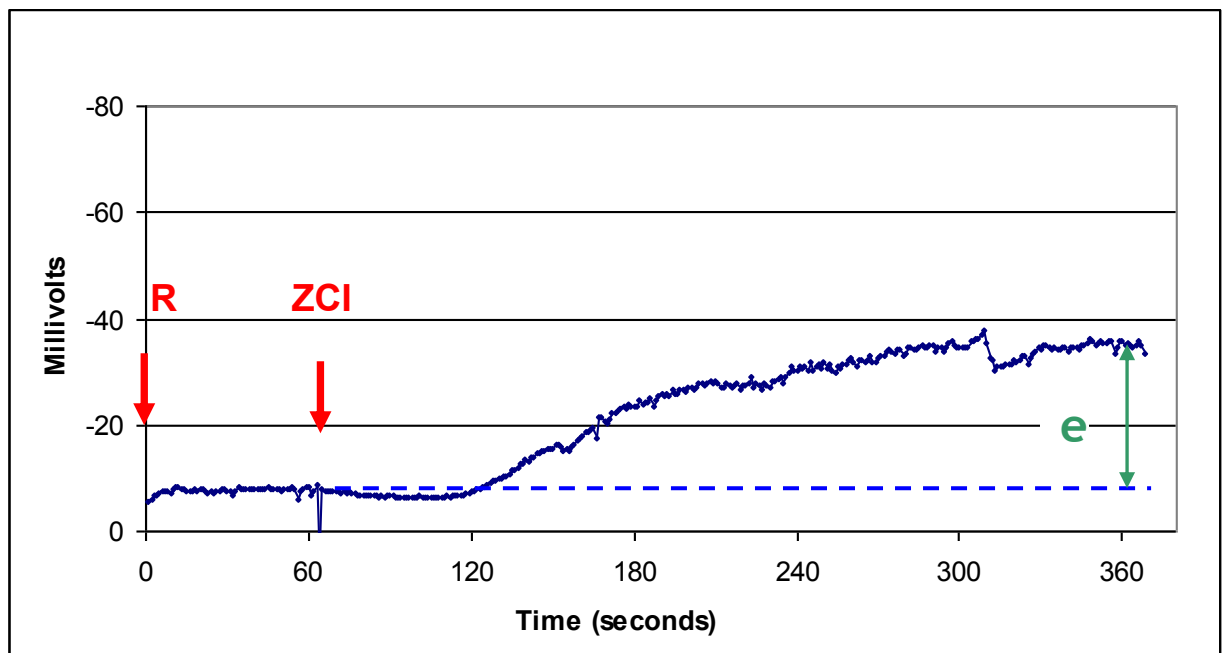


Figure 2.16. Initial bronchial PD set up tested in the non-CF nose. The lower airway PD catheter was applied to the nasal epithelium. Nasal PD was measured (indicated by navy coloured line) during perfusion with Ringers ('R') and then with Zero Chloride isoprenaline solution (ZCI; the start of ZCI perfusion is marked with the corresponding arrow) using the initial bronchial PD set up. The light blue dashed line represents the basal PD (in Millivolts) during perfusion with Ringers solution. The response to ZCI solution after 5 minutes is indicated by the change in PD (delta) between start and end of ZCI perfusion, indicated by 'e'.

This test set up demonstrated ability to achieve stable basal measurements with Ringers solution and to assess change in PD as a consequence of Cl^- secretion in response to the zero chloride solution perfusion. This system was then used for lower airway PD measurements in the Pilot study.

All patients undergoing bronchoscopy were in Pilot part A, as detailed in section 2.2. To briefly summarise, all patients received a 20ml nebulised lung dose and all had both nasal and lower airway PD measurements performed at time points as defined in the trial protocol. Bronchoscopies were performed by JD (n=21) and AB (n=3) with bronchial PDs performed by GD (n=21) and JD (n=3). All bronchoscopies were done under general anaesthetic as per trial protocol with written informed consent obtained pre-procedure.

Skin PD was measured using the tubing set up primed to the tip with Ringers solution once the patient was anaesthetised to check integrity of the bronchial PD measuring 'circuit'. The cytology brush sheath re-primed with Ringers was then inserted into the bronchoscope in preparation for lower airway measurements.

Basal bronchial PD measurements were made after a minimum of 20 seconds stability during perfusion with Ringers solution at 100ml/hr at both lateral surfaces, anterior and posterior surfaces of the main carina. Basal PD was then measured in the first of three distal (approximate 5th generation bronchi) sites. Once a stable trace had been achieved, perfusion was switched to zero chloride isoprenaline solution at 100ml/hr for five minutes. The sheath was then removed from the bronchoscope and the circuit purged with 2.5mls of Ringers solution to eliminate the column of zero chloride solution; second and third distal basal and perfusion measurements were performed in the same way. All distal measurements were performed on the same side. The decision to make measurements in either the left or right lung was based largely on the macroscopic appearances including presence of secretions which could potentially interfere with measurement.

2.6.3.1 New problems encountered

Two patients underwent bronchoscopy with the initial set up described above. However, several problems were identified which were not apparent from the initial tests carried out in healthy volunteers using the nasal epithelium as described.

1. The circuit set up for both these patients was characterised by high offsets (-27.6 and -23.0 mV) and difficulty in obtaining accurate traces due to the

frequent intermittent loss of electrical contact and an inability to achieve a stable basal measurement.

2. Visible small bubbles were trapped within the circuit soon after drawing up the solutions.
3. Loss of electrical contact occurred during PD perfusions as a result of bubbles within the circuit.

It became apparent that additional bubbles were also forming *during* the PD measurements; these could not have been prevented as they were not present at the time of circuit preparation.

The author hypothesized that the appearance of new bubbles related to the rising temperature of the refrigerated solutions. To investigate this further, syringes and tubing from the second bronchoscopy were kept at room temperature until the following day by which point thousands of tiny bubbles were distributed throughout the tubing and lining the syringe walls. The author reasoned that such bubbles may not form if the solutions were warmed prior to being drawn up into syringes and then allowed to cool to room temperature. This hypothesis was confirmed in the laboratory: no bubbles formed within tubing or syringes if the PD tubing circuit was prepared with pre-warmed solutions.

The three-way taps within the tubing circuit also appeared to contain air bubbles when switching between open ports. Their opaque centre meant that it was not possible to visualise the entire fluid columns within the circuit.

2.6.4 Further development and use in Pilot study

Further modification of the circuit was therefore undertaken (Figure 2.17) to tackle high offsets and bubble formation:

1. Solution warming: Solutions were warmed in the RBH operating theatres' fluid warmer to approximately 37 degrees C (stock PD solutions were removed from the refrigerator and placed in the warmer on the afternoon before the test).

2. Slow and careful drawing up of solutions into 50 ml syringes from EVA stock PD solution bags.
3. Minimising connections: All three-way taps were removed from the circuit. Instead, T-piece connectors with rubber bungs (Medex, Smiths Medical) allowed a '3-way' connection by a primed needle piercing the bung and clamping until required. This preserved the ability to perfuse sequential solutions independently.

A diagram summarising the modified circuit is shown in Figure 2.17.

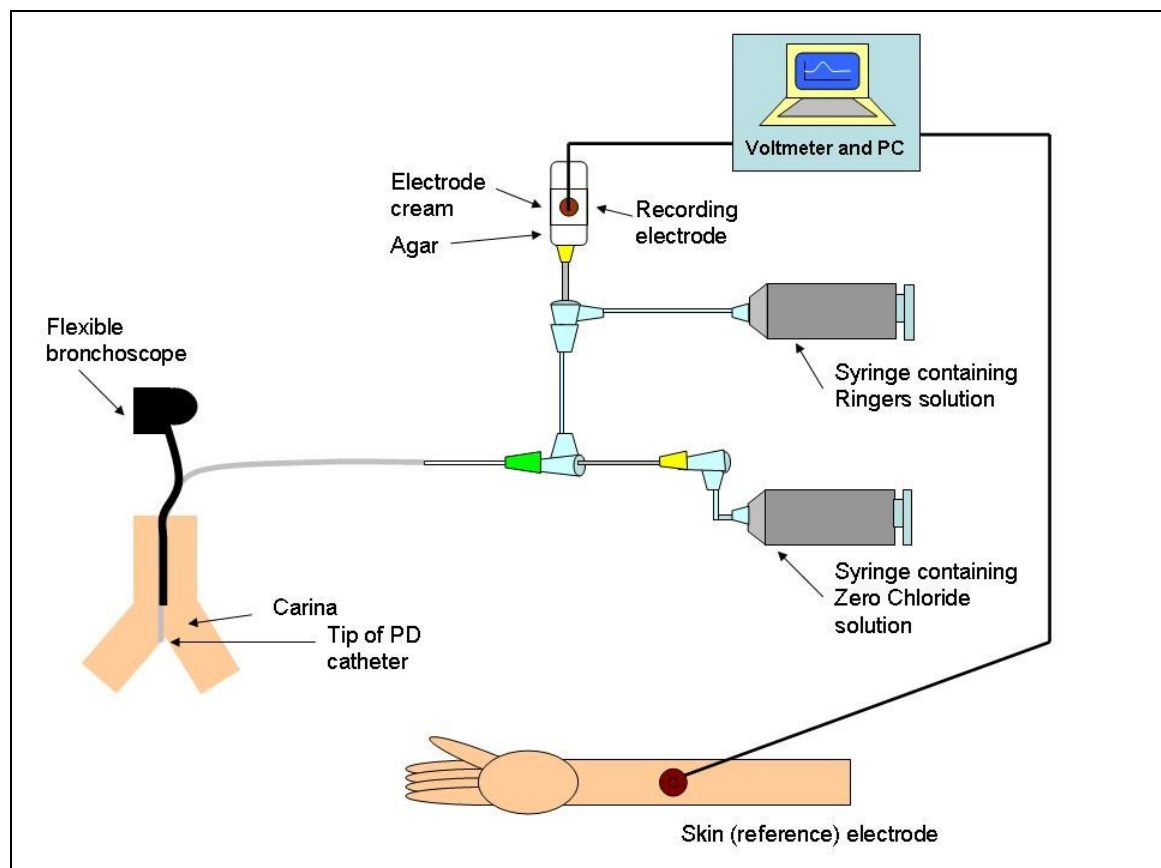


Figure 2.17. Schematic representation of the modified bronchial PD set up. All solutions were warmed prior to circuit preparation. PD catheter (cytology brush sheath) is connected to t-piece connector via an Abbocath catheter. A needle pierced through the t-piece connector bung is used as a three way connection to allow perfusion with zero chloride isoprenaline solution once baseline Ringers measurements have been performed (no 3-way taps are in the circuit). The transition between Ringers and Zero chloride isoprenaline solutions during PD measurements is made by clamping off the relevant t-piece connector.

This modified circuit was used to perform all PD measurements in patients (n=10) with pre- and post gene therapy bronchoscopies, and a further two patients with a single pre- bronchoscopy who withdrew from the Pilot study prior to dosing. Offsets

with the modified circuit were substantially lower than those prior to circuit optimisation: the mean (SD) offset for pre-dosing tests was -3.2 (2.1) mV and -3.5 (2.0) mV for post dosing tests. The Offsets were corrected to zero prior to making measurements of skin PD and through the bronchoscope.

All patients with bronchial PDs undertaken with the modified circuit had carinal and distal measurements made successfully. On occasion, a stable distal baseline was not achieved. Where this occurred, that site was abandoned and an additional site tested so that all patients had three perfusion results. Taking three distal measurements each with a subsequent perfusion stage enabled the calculation of a delta to quantify chloride secretion at each site which could be averaged to take account of variability.

2.6.5 Data recording and analysis in Pilot study

A proforma was used to record the timing of all basal measurements and start and end of perfusion solutions (in seconds post start of data recording). This was crucial to guide the subsequent analysis of results. Timing of basal and perfusion measurements was documented at each site.

Data files for each test were saved according to patient code and PRE or POST denoting status in relating to gene therapy dose. Creation of the .txt files was undertaken as per nasal PD (Chapter 2, section 2.5).

Data analysis of all traces was undertaken following completion of Pilot study data collection. The method of creating a display of data in Excel was identical to that described for nasal PD in (Chapter 2, section 2.5).

For blinded data analysis carried out by the author, filenames identifiable with patient code were replaced with a 3 digit random number (generated with Excel random number generator). All identifying features including patient ID and relationship to dosing (pre/post) were removed. Analysis was then performed using the time points recorded in real time during bronchoscopy for each site and start/end of perfusion on anonymised bronchial PD proformas.

To identify the PD at each point of interest (at the designated time points for measurement (as defined on the proforma) and corresponding visual analysis of the trace (displayed graphically in Excel)), a mean of up to 20 values (20 seconds) of mV recording were used to identify the mV result at each time point. The number of time points averaged by the author depended on stability of the trace at the specified time point. For example, where an equal (bidirectional) variation in mV was seen n=20 time points were used, whereas if the trace appeared to be on an incline a smaller number of data points (<20) were averaged according to visual perception of 'true' mV value at that point. Use of mean values calculated electronically rather than isolated single mV (single second) readings were used to improve accuracy of the PD.

For each analyser (the author (GD) and JD), mean basal carinal measurement (average of anterior, posterior, right and left lateral surfaces), mean distal basal PD (average of the three measurements), and mean delta chloride response (average of the three deltas) were calculated.

The data from all patients with paired pre- and post bronchial PD measurements were analysed independently by the author and JD prior to combining analyses and consensing on final PD values for each PD variable of interest. This is described in more detail below.

2.6.5.1 Combining two lower airway PD analyser results

The final agreed mV value for each parameter was determined as follows:

- Difference between GD and JD mV values for mean carinal basal, mean distal basal and mean of the n=3 distal delta chloride responses was calculated
- Values for each analyser were automatically accepted (and not other technical concerns about the trace) if the difference between GD and JD was within following pre-defined range:
 - Mean carinal basal measurement: 2mV maximum difference allowed

- Mean distal basal measurement: 2mV maximum difference allowed
- Mean Zero chloride isoprenaline delta: 1mV maximum difference allowed
- Tests where difference between GD and JD was greater than these allowed values were discussed on a test-by-test basis with visual analysis of the trace (still anonymised and blinded) in combination with the time points recorded on proforma during bronchoscopy. If, following review of plotted values, both JD and GD were considered equally valid, both were accepted and the mean calculated.
- Tests deemed non-analysable by both analysers were excluded from final data analysis
- Where disagreement out with the above criteria existed, PD traces were reviewed by JD/GD and a consensus was agreed about final values for data analysis.

2.6.5.2 Final lower airway PD values

All results presented here are from the final GD/JD combined values.

Although post dosing bronchoscopies were performed at one of two different time points post dosing (day 6 or 14), to allow comparison between pre- and post-, measurements bronchoscopies performed on either date were combined for data analysis.

2.6.5.3 Analysis of final lower airway PD and comparison with other outcome measures

The following analyses and comparisons were undertaken using the final lower airway PD results including:

- Comparison between basal measurements obtained proximally (carina) and distally

- Change in mean basal (mV) PD measurements between pre and post dosing at proximal and distal sites.
- Change in response to zero chloride/ isoprenaline perfusion between pre- and post dosing. This was analysed on a within patient and between group (pre vs. post dose bronchoscopy) basis.
- Comparison of bronchial PD with nasal PD pre- and post dose: accepting differences between sites (drug delivery, disease manifestation and structure), this was undertaken to determine whether there is any relationship in terms of ion transport and whether one can be used as a surrogate for the other.

2.6.6 Statistics

2.6.6.1 Agreement between analysers

Intraclass correlation coefficients (ICCs) were calculated to assess agreement on the lower airway PD values as calculated by GD and JD using SPSS software as described in Chapter 2, section 2.5. This was undertaken in two phases, using the same approach as nasal PD:

- Initial agreement between all lower airway PD analysed values
- Final agreement between acceptable tests following consensus opinion between GD and JD.

ICCs were calculated using SPSS Statistics version 21.0 (IBM), using a two way mixed effects model for absolute agreement. Final PD values for each section of PD (mean of accepted GD and JD result) were combined to analyse group differences between pre- and post- measurements. Comparison between groups was performed using paired t-tests with significance at $p < 0.05$. Pearson correlation coefficients were used to describe associations between continuous variables.

2.7 Functional efficacy outcome measures: molecular evidence of *CFTR* gene transfer

In the Pilot study, in addition to airway PD measurements, molecular evidence of successful *CFTR* gene transfer was also investigated. This was undertaken in the upper and lower airway for patients receiving a nasal or lung dose of gene therapy respectively.

2.7.1 Nasal brushing for molecular *CFTR* assays

Nasal brushings were performed pre- and post, or post dosing only for subjects receiving a nasal dose of CF gene therapy. All nasal brushings were performed by JD. Interdental brushes (Dent-o-care, London, UK) were used for sample collection dipped in phosphate buffered saline (PBS) first, with samples taken from under the middle or inferior turbinate under direct vision and immediately placed in 800 microlitres PBS. Cells were removed from the brush by passing it back and forth 20 times through a wide pipette tip into the PBS. A sample was taken from both nostrils, and combined, before being split three ways for subsequent analysis for immunocytochemistry, percent positive cells (PPC) and RNA extraction. Cells obtained from nasal brushing were counted using light microscopy for differential cell numbers. Molecular testing was carried out by Consortium researchers in Oxford for detection of vector-specific *CFTR* mRNA and DNA. Percent positive cells were analysed using the Taqman^(R) Gene Expression Cells-to-CTTM Kit.

2.7.2 Bronchial brushing for molecular *CFTR* assays

Samples for molecular *CFTR* testing (DNA and mRNA) were obtained from lower airway brushings obtained at bronchoscopy. All bronchoscopies were performed at the RBH under general anaesthetic. The sequence of tests performed whilst a patient was anaesthetised included: lower airway potential difference measurements, bronchial brushings, and endobronchial biopsy. Bronchial brushings were processed and tested as described for nasal brushings above. Bronchial brushings were taken using disposable cytology brushes (BC-202D-5010, Olympus UK, Southend on Sea, UK). Ten brushings were taken per patient per bronchoscopy.

In addition to the bronchial brushing samples which were taken for the purposes of assessing efficacy, two endobronchial biopsies were taken per patient at each bronchoscopy using disposable single-use forceps (FB-231D, Olympus UK, as before). Biopsies were immediately fixed in formaldehyde. Histological analysis for presence of inflammation or evidence of remodelling was carried out by the histopathology department at the RBH according to standard hospital procedures for analysis of endobronchial specimens.

Where time allowed for same-visit nasal PD to be performed pre-bronchoscopy, nasal brushing was also performed under anaesthetic to minimise discomfort for the patient. Patients were observed and monitored as in patients at RBH for a minimum of six hours post bronchoscopy.

3 Results: Run-In study

This chapter describes the results from the first four Run-In study visits. This includes description of the study population and results (cross-sectional and longitudinal) for airway physiology, structural, patient reported outcomes and systemic markers of inflammation.

3.1 Patient recruitment and characteristics of study population

192 patients were recruited to the Run-In study, with 191 patients completing 1 or more study visits (n=111 at London site, n= 80 Edinburgh). Figure 3.1 shows the number of participants over the duration of the 4 visits.

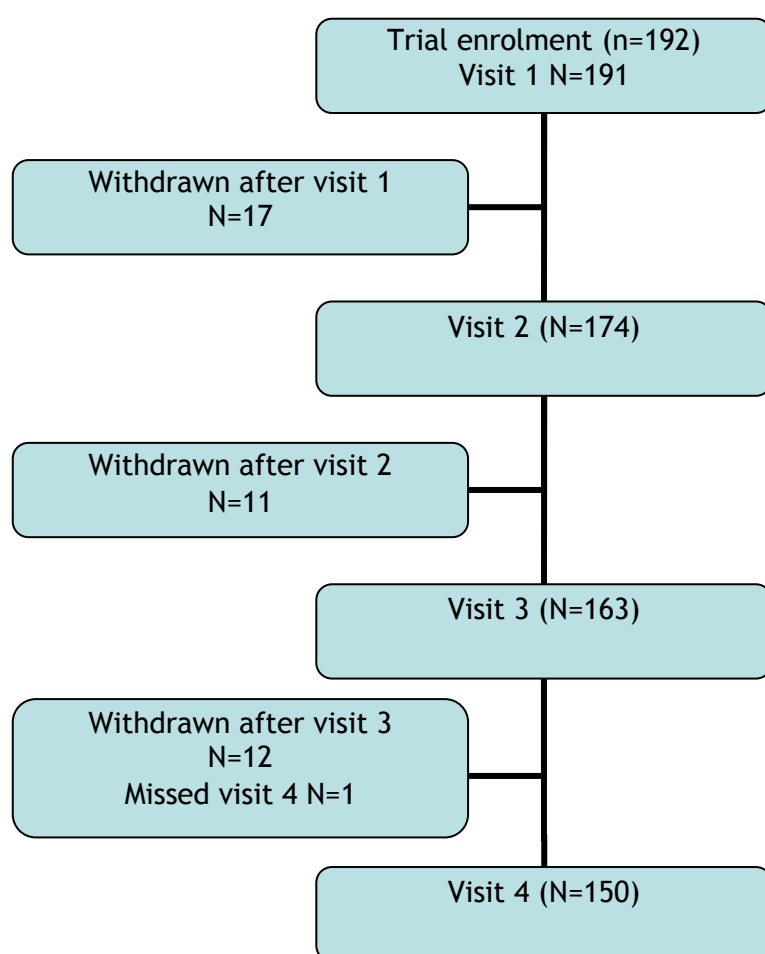


Figure 3.1. Number of patients (N) participating in each Run-In study visit.

The subject attrition rate between visits 1-4 was linear (the average decline between visits was n=13 subjects, or 7% loss in study population; linear regression $r^2 = 0.99$).

3.1.1 Subject characteristics

At visit 1, the median (IQR) age was 18.7 (14.0-28.7). Ninety one subjects were aged <18 years at enrolment (London n=50, Edinburgh n=41), with the youngest patient aged 10 years. The oldest patient recruited was 59 years. Other baseline characteristics of Run-In patients are summarised in Table 3.1.

		N (%)
Gender	Male	105 (55)
	Female	86 (45)
Pancreatic insufficient?	Yes	167 (87.4)
	No	24 (12.6)
Diabetes mellitus	Yes	36 (18.8)
	No	155 (81.2)
IV antibiotics for exacerbation in past 12 months?	Yes	73 (38.2)
	No	118 (61.8)
Regular elective IV antibiotics?	Yes	20 (10.5)
	No	168 (88.0)
	Unknown	3 (1.6)
Airway pathogens identified in past 24 months	<i>P.aeruginosa</i> (any)	121 (63.3)
	<i>P.aeruginosa</i> (mucoid)	77 (40.3)
	<i>S.aureus</i>	88 (46.1)
	<i>H.influenzae</i>	31 (16.2)
	<i>S.maltophilia</i>	22 (11.5)
	<i>A.fumigatus</i>	53 (27.7)

Table 3.1. Baseline characteristics of Run-In study participants

The median FEV₁ % predicted for the group as a whole at visit 1 was 71.0 % (Interquartile range (IQR) 53.5-84.8), which corresponded to a median FEV₁ z score of -2.5. Patients recruited at the London site had more severe disease (as assessed by FEV₁) than those in Edinburgh (median (IQR) 63.1 (48.5-78.2) % and 78.4 (63.4-

90.4) % respectively, Mann Whitney test $p < 0.0001$). The patients recruited at the London site appeared to be older however there was no statistically significant difference between their age and those of the Edinburgh patients at visit 1 (Mann Whitney test $p = 0.08$). There was no difference in gender split between sites (52% male in London, 60% male in Edinburgh; Chi square test $p = 0.2$).

Almost fifty percent of the study patients were homozygous for the p.Phe508del mutation. Genotypes from all patients in the Run-In study are summarised in Table 3.2.

	p.Phe508del	Other	Unknown
p.Phe508del	95 (49.7)	62 (32.5)	20 (10.5)
Other		9 (4.7)	4 (2.1)
Unknown			1 (0.5)

Table 3.2. Genotypes from Run-In study participants. CF mutations are represented in grid format, mutation combinations in study patients are summarised according to number of patients (N (%)) with each combination.

3.1.2 Interval between visits

The number of days between study visits is summarised in Figure 3.2. The interval between visits 1 and 2 was significantly shorter than between visits 3 and 4 (median 113 vs. 182 days; Wilcoxon test $p < 0.0001$). This pattern was observed in both London and Edinburgh sites and reasons are described in the Discussion (section 3.7.2).

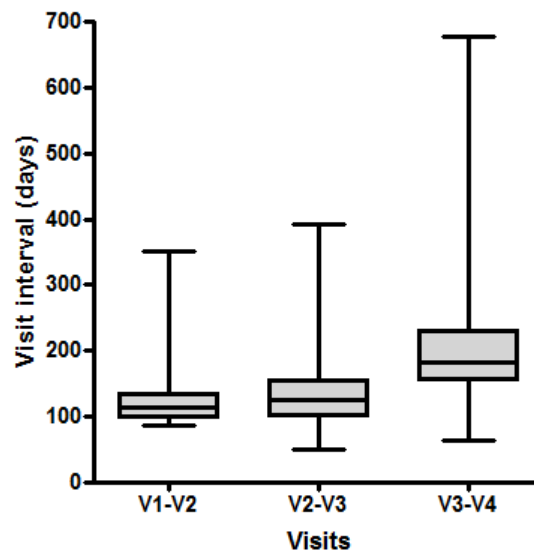


Figure 3.2. Interval between study visits 1-4. Results are summarised as Box and Whisker plots.

3.2 Description of data

The Run-In data provide opportunities to a) look at the variability of assays over time during periods of stability; b) question how these assays correlate with others, both within and across categories (e.g. physiology, inflammation, radiology), and c) observe directional change over time. Sections 3.3-3.5 focus on the first two of these as ‘interval independent’ analyses. A subsequent section (section 3.6) describes changes over time. For consistency, the data in section 3.3 are presented for those patients with results for a particular assay at every one of the 4 study visits. The assays described and the number of patients with all 4 data values available are summarised in Table 3.3. For many of the analyses described in this chapter, visit 1 data are used, whereas for others, particularly when looking at repeatability, the group of patients with data from all 4 study visits was used. To confirm whether these two groups were similar in baseline characteristics, Table 3.3 includes visit 1 data from both groups. With the exception of LCI, there was no overall difference between subjects with results at every one of the study visits and those in the entire study population (Table 3.3); the potential reasons for this difference are described in the Discussion (section 3.7.4).

	Total No. patients with valid assay results for all 4 visits	Mean (SD) assay result at V1	Total No. patients with valid assay result at V1	Mean (SD) assay result at V1	Comparison between groups
Physiology					
FEV ₁ z score	150	-2.45 (1.67)	191	-2.57 (1.62)	p=0.51
FVC z score	150	-1.51 (1.52)	191	-1.59 (1.45)	p=0.64
MEF ₂₅₋₇₅ z score	150	-2.60 (1.76)	191	-2.72 (1.71)	p=0.52
LCI	92	10.03 (2.57)	168	10.75 (2.73)	p=0.04
Inflammation					
Total white cell count	137	9.16 (2.86) x 10 ⁹ /L	187	9.32 (2.96) x 10 ⁹ /L	p=0.63
Blood Neutrophils	137	6.05 (2.64) x 10 ⁹ /L	186	6.20 (2.74) x 10 ⁹ /L	p=0.65
CRP	137	6.40 (10.91) mg/L	186	6.55 (10.15) mg/L	p=0.90
Quality of life					
CFQ-R Respiratory domain	144	78.61 (14.93)	190	77.44 (15.64)	p=0.49

Table 3.3. Comparison of visit 1 outcomes from patients with assay result at every one of the 4 study visits with the initial study population. The visit 1 (V1) outcomes from patients with valid results for an assay at all 4 study visits are compared to the entire study population with a valid assay result at visit 1. FEV₁, FVC and MEF₂₅₋₇₅ values are for z scores. Unpaired t-tests were used to compare V1 results for patients with data from all 4 visits to the entire study population, with p values as shown.

In the following sections, I will present i) a summary of the results for each outcome variable at the first study visit and its subsequent variability across visits for patients with data from all 4 visits available (section 3.3), ii) cross-sectional correlations within category (e.g. physiological outcomes) (section 3.4) and iii) cross-sectional correlations between categories (e.g. FEV₁ z score with Respiratory domain of CFQ-R) (section 3.5). For the purposes of consistency, as a CT was performed only at one visit, all cross-sectional analyses are performed on data from that visit.

3.3 Summary data and variability

3.3.1 Physiology: Spirometric values

A Box and Whisker summary plot of the % predicted values and FEV₁ z scores at each study visit from all patients with data for all 4 visits are shown in Figure 3.3.

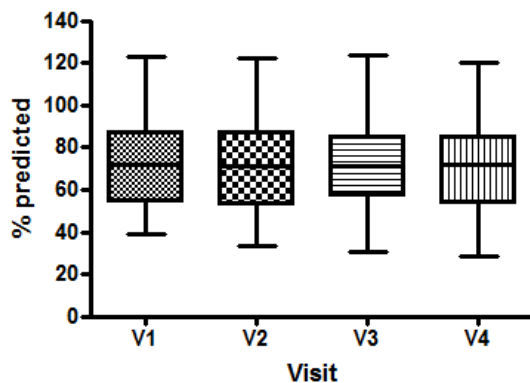


Figure 3.3A

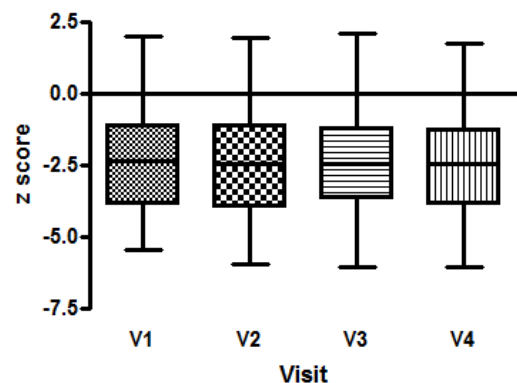


Figure 3.3B

Figure 3.3. FEV₁ % predicted values and z scores across all 4 study visits. FEV₁ % predicted values are shown in 3.3A and z scores in 3.3B. Data are summarised in Box and Whisker plots.

The median inter-visit coefficient of variation in FEV₁% predicted values was 6.04%. Patients with an FEV₁% predicted in the lowest quartile (V1 FEV₁ < 54.5%) had greater variability than those in the highest quartile (V1 FEV₁ ≥ 87.1%), with a median coefficient of variation of 7.83% vs. 5.10% respectively, Mann Whitney p=0.02). Using a cut off of -1.64 z scores for FEV₁ as the lower limit of normal (LLN) (116;117) (approximately equivalent to an FEV₁ of 80% predicted), Figure 3.4 illustrates the inter-visit coefficient of variation for patients with results at each visit. Patients with a normal FEV₁ z score at visit 1 appeared to show a lower variability in results but this was not statistically significant.

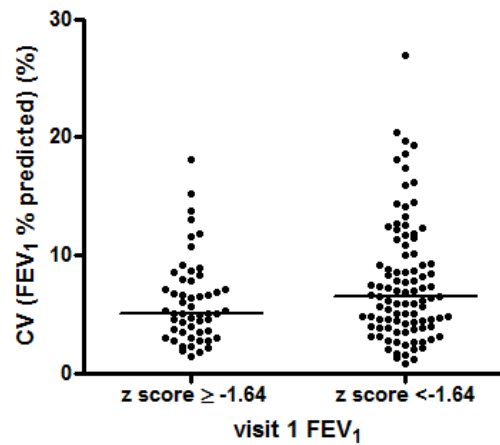


Figure 3.4. Coefficient of variation in FEV₁% predicted according to visit 1 FEV₁ z score. Each patient is represented by an individual data point. A z score of -1.64 was the cut off between normal and abnormal lung function at visit 1; coefficient of variation (CV) was plotted for each group. Mann Whitney test p=0.1.

Using the same format as for FEV₁, the results for FVC and MEF₂₅₋₇₅ % predicted and z scores, and ratio of FEV₁/FVC are shown in Figure 3.5. The median inter-visit coefficient of variation in MEF₂₅₋₇₅ predicted values was 13.97%. Patients with a higher inter-visit coefficient of variation in FEV₁ % predicted also had a higher coefficient of variation in MEF₂₅₋₇₅ % predicted ($r = 0.73$, $p < 0.0001$). For a patient with an inter-visit FEV₁ % predicted coefficient of variation of 'x', their corresponding coefficient of variation for MEF₂₅₋₇₅ % predicted was approximately '2x' (linear regression analysis: slope 2.10, 95% CI for slope 1.96-2.24). In other words, a patient with a 10% coefficient of variation for FEV₁% predicted was likely to have had a 20% coefficient of variation in MEF₂₅₋₇₅ % predicted values across the 4 study visits.

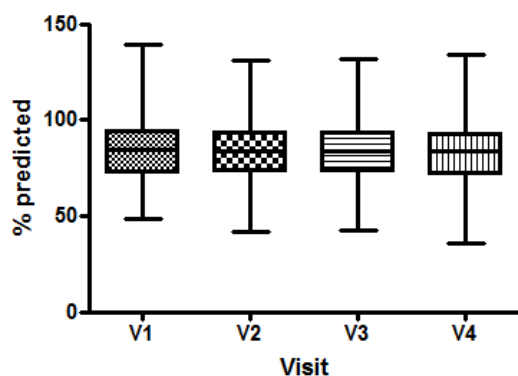


Figure 3.5A % Predicted FVC

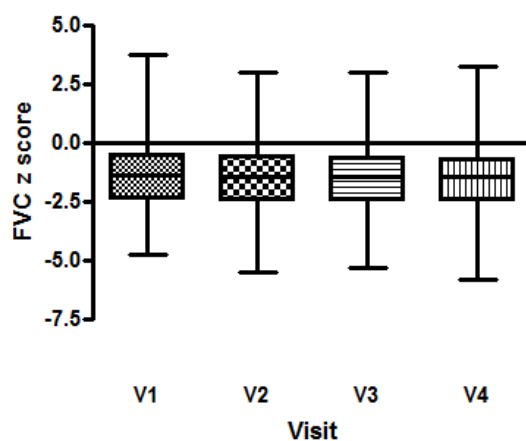


Figure 3.5B FVC z score

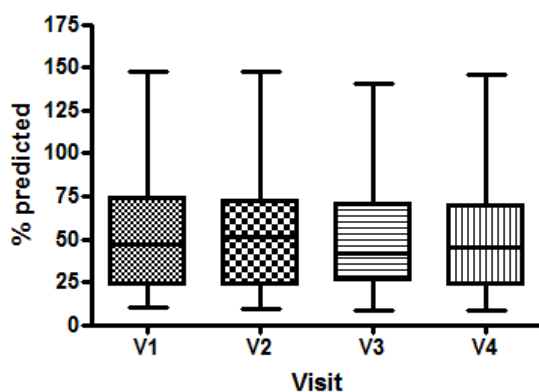


Figure 3.5C MEF₂₅₇₅ % predicted

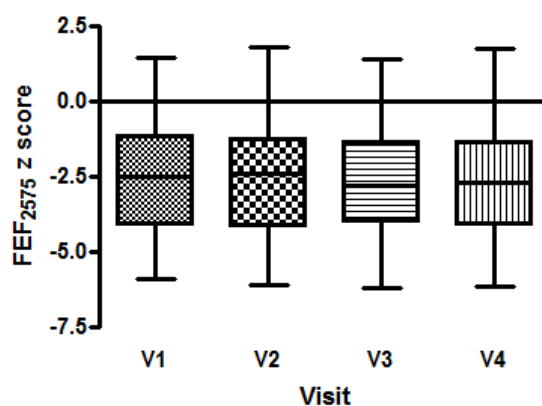


Figure 3.5D MEF₂₅₇₅ z score

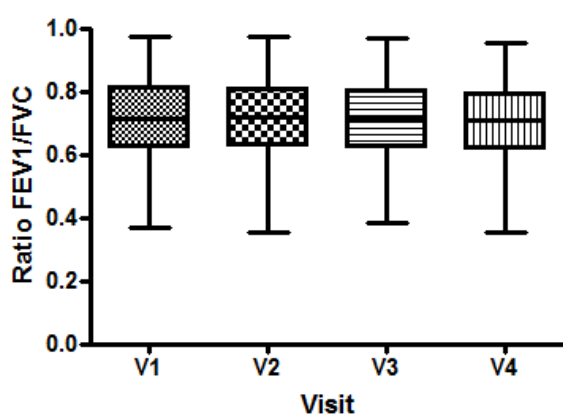


Figure 3.5E Ratio of FEV₁/FVC

Figure 3.5. Summary data for spirometry outcomes at each study visit. Data for all 4 visits summarised as previously described for the outcome indicated on graphs A-E. The summary data for FEV₁ is shown in Figure 3.3.

3.3.2 Physiology: Lung clearance index

The summary final LCI results are summarised in Figure 3.6. The dataset was less complete for LCI than other outcome measures, reflecting the processes in data analysis and quality control described in Chapter 2, section 2.3.

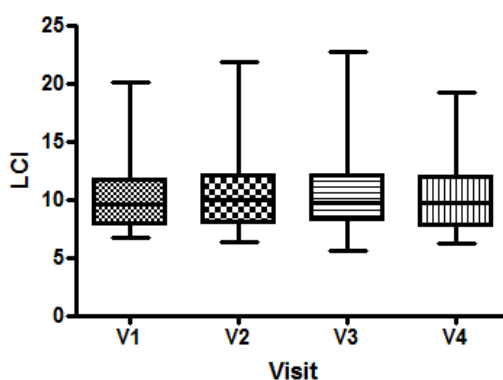


Figure 3.6. Summary data for lung clearance index (LCI) at each study visit.

Only seventeen patients had a normal LCI at visit 1 according to the Upper limit of normal (ULN) as defined by Horsley (ULN 7.5 in adults and 7.3 in children ≤ 16 years) (48). The median inter visit coefficient of variation in LCI was 6.4%. The relationship between severity and variability was explored. Patients in the top (most severe) quartile for visit 1 LCI had greater variability than those in the bottom (mildest) quartile (Mann Whitney $p=0.03$) (Figure 3.7). Similarly, those with a normal LCI at visit 1 had less variability over subsequent visits than those with an abnormal LCI (Mann Whitney $p=0.04$) (data not shown), although for the group as a whole, there was no correlation between baseline (visit 1) LCI and coefficient of variation across the 4 visits ($r=0.10$, $p=0.4$).

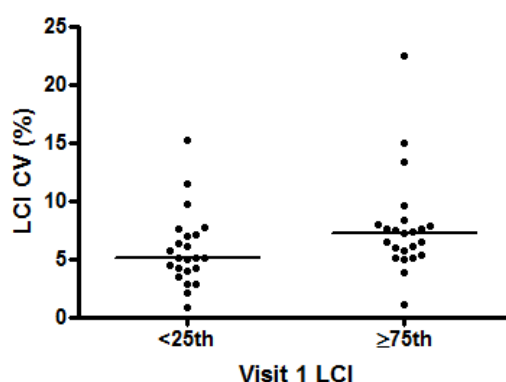


Figure 3.7. Variability in LCI between study visits according to quartile of visit 1 LCI in Run-In population. The inter-visit coefficient of variation (CV) for each patient in the lowest (<25th) and highest (≥75th) quartile is represented by an individual data point. Mann Whitney test p=0.03.

3.3.3 Structure

180 baseline chest CT scans were performed for the Run-In study. Of these, 178 have been formally scored (data not available for 2 patients). Scores are available on 131 scans performed at visit 1, 45 at visit 2, and two at visit 3. Eleven patients had not undergone a CT chest scan according to the CTEU database: 7 patients in this group withdrew after visit 1 and therefore did not have opportunity for CT at later visits. The reason for the absence of CT data for the other 4 subjects is unknown.

The Run-In CT data are summarised in Table 3.4. Only 27/178 (15.2%) patients had no evidence of bronchiectasis on their CT scan (CT score <1 for extent of bronchiectasis). These patients tended to be young (median (IQR) age 13.4 years (11.0-17.7)), had mild disease according to spirometry measurements (median (IQR) FEV₁% predicted 89.3 (81.9-96.1)), and have normal or near normal LCIs (median (IQR) 7.5 (6.9-8.6)) at visit 1.

	Median	IQR	Range	Maximum (ceiling) score
Extent of bronchiectasis	2.7	1.5-3.4	0-4	4
Severity of bronchiectasis	2.0	1.2-2.3	0-3.5	4
Wall thickening	2.0	1.6-2.1	0.9-3.8	5
Small airway plugs	0.6	0.2-0.8	0.0-1.8	2
Large airway plugs	0.6	0.2-1.0	0.0-1.8	2
Air trapping (%)	59.6	54.6-65.4	31.7-90.8	100

Table 3.4. Summary results from chest CT scans. Summary results from n=178 patients at a single time point scored according to the methodology described in Chapter 2, section 2.4.3.

3.3.4 Inflammation

The results for total white cell count, neutrophil count and C-reactive protein for patients with results at 4 visits are summarised in Figure 3.8. The median inter visit coefficients of variation for these measurements for all patients with results at all 4 visits are summarised in Table 3.5 (section 3.3.6). The increased variability of CRP compared to white cell count likely reflects the very small normal range (<10 mg/L) and the potential for very high values in inflammatory conditions. A small number of patients had high CRP values despite the presence of clinical stability as assessed by symptom questionnaire performed at study visit. The 75th centile for CRP was normal (<10mg/l) and consistent between study visits (7- 8 mg/L) however the maximal result for CRP (in any study patient) decreased between visit 1 and 4 from 109 to 38mg/L.

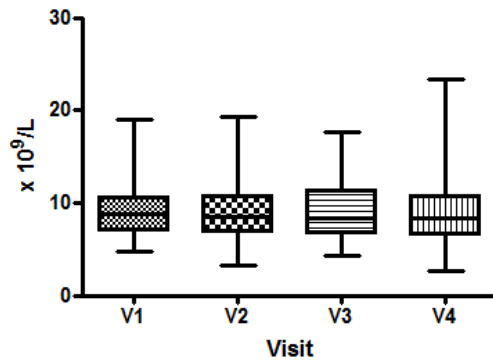


Figure 3.8A Total white cell count

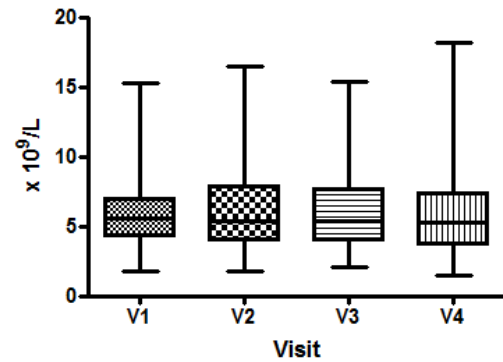


Figure 3.8B Neutrophil count

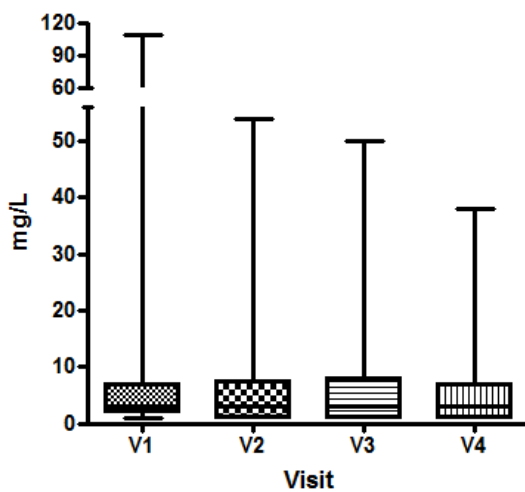


Figure 3.8C C-reactive protein

Figure 3.8. Blood markers of systemic inflammation. Summary plots are shown for blood total white cell count (Figure 3.8A), neutrophil count (Figure 3.8B) and C-reactive protein (Figure 3.8C).

3.3.5 Quality of life questionnaires

Results were combined for scores from adult and paediatric questionnaires, with one questionnaire per visit. 141 adult and 49 paediatric questionnaires were completed at visit 1 (results from 1 paediatric patient are unknown). The summary results from all study patients each questionnaire domain at visit 1 is shown in Figure 3.9. The median score was greater than 90 for the following domains at visit 1: Physical, Role, Eating disturbances, Weight (the maximum score possible for any domain is 100).

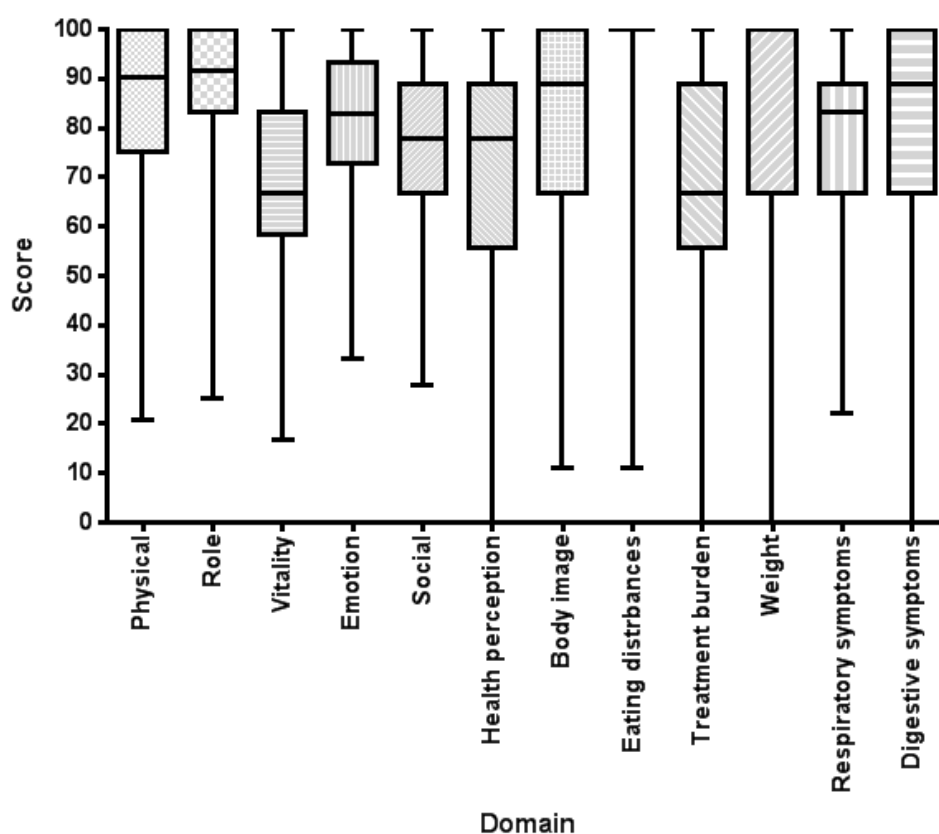


Figure 3.9. CFQ-R scores for each questionnaire domain at Run-In visit 1.

The ‘Respiratory symptoms’ domain scores at each study visit are summarised in Figure 3.10; the median inter-visit coefficient of variation for this domain was 8.7%.

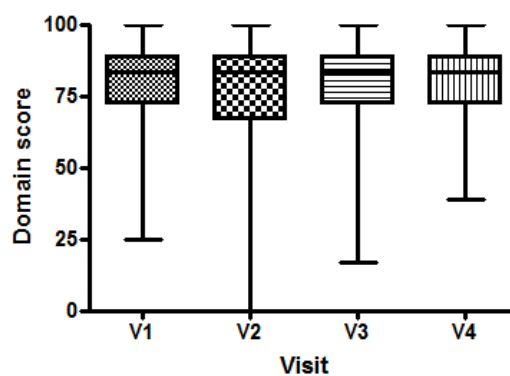


Figure 3.10. CFQ-R Respiratory domain score summary data at each study visit.

Nine patients transitioned between the paediatric and the adolescent/adult questionnaire over the study period. There was no clear directional bias to change in score between visits when transitioning between paediatric and adult questionnaire versions. The median inter-visit coefficient of variation for these adolescents across all 4 visits (i.e. including responses from both paediatric and adult questionnaires) was 9.8%. The variability in inter-visit Respiratory symptom domain score from the CFQ-R was significantly greater in those patients with an abnormal FEV₁ z score at visit 1 (Mann Whitney test p= 0.03).

3.3.6 Summary

In summary, this section has described the results and variability over time for the physiology, inflammatory and patient reported outcomes described in this chapter. For comparison in variability between outcomes, the median inter-visit coefficient of variation for each is summarised in Table 3.5.

Outcome measure	Median CV (%)
Physiology	
FEV ₁ %	6.0
FEV ₁ z score	N/A
FVC %	4.5
FVC z score	N/A
FEV ₁ /FVC	3.5
MEF ₂₅₋₇₅ %	14.0
MEF ₂₅₋₇₅ z score	N/A
LCI	6.4
Inflammatory	
Total white cell count	14.8
Blood neutrophils	21.7
CRP	31.6
Quality of Life	
CFQ-R Respiratory domain	8.7

Table 3.5. Mean inter-visit coefficient of variation (CV) between visits for the outcome of interest. CV was not calculated for z scores due to the proximity to zero and presence of both positive and negative values.

3.4 Cross-sectional correlations within categories

To allow a greater understanding of the clinical importance of the outcomes investigated, the relationships between them has been explored. This section describes cross-sectional correlations of outcomes within each category (e.g. airway physiology). These relationships help understand whether outcomes reflect similar or ‘additional’ information. The latter may either be due to increased sensitivity to detect disease or differences in site of airway under test, allowing a clearer understanding of disease pathology and which level of airways the different assays are representing. As patients only had one CT, all measurements included in this analysis are from that visit (1st, 2nd or 3rd as described above).

3.4.1 *Airway physiology*

Correlations between spirometry measurements (FEV_1 , FVC and MEF_{25-75}) and LCI (%) and z score) are summarised in Table 3.6 for measurements performed at the CT visit. The strength of correlation between FEV_1 , FEV_1/FVC and MEF_{25-75} with LCI were virtually identical despite FEV_1/FVC and MEF_{25-75} considered as reflecting different sites of the airway (Table 3.6 and Figure 3.11) (larger vs. small airways). To investigate sensitivity, the results for FEV_1 z score vs. LCI and MEF_{25-75} vs. LCI have been plotted with reference to a line separating normal and abnormal measurements (Figure 3.11). For both FEV_1 z score and MEF_{25-75} , LCI was more sensitive outcome measure. Conversely the number of patients with an abnormal z score but normal LCI represented less than 3% patients.

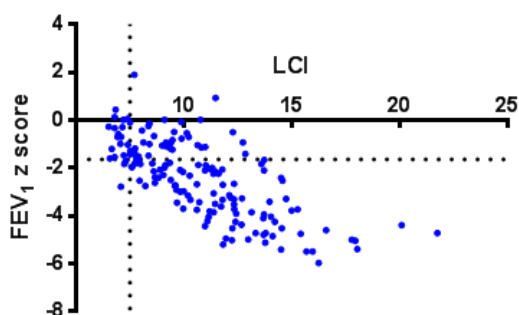


Figure 3.11A

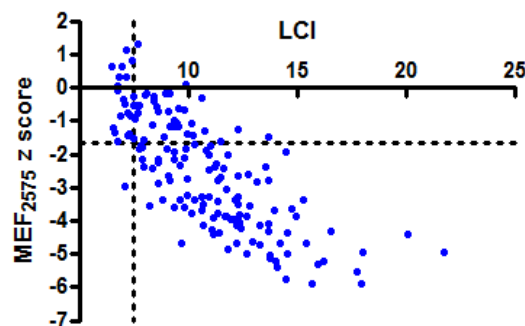


Figure 3.11B

Figure 3.11. Spirometry and LCI correlations. Each patient is represented by a single data point. Dotted lines separate normal from abnormal results. This is plotted at -1.64 z scores for FEV_1 (Figure 3.11A) and MEF_{25-75} (Figure 3.11B) and at 7.5 to differentiate normal from abnormal LCI in both figures. Abnormal results are defined as FEV_1 or $MEF_{25-75} < -1.64$ z scores or an LCI > 7.5 .

	FVC z score			FEV ₁ /FVC			MEF25-75 z score			LCI		
FEV ₁ z score	180	0.86	<0.0001	180	0.76	<0.0001	180	0.91	<0.0001	156	-0.72	<0.0001
FVC z score				180	0.35	<0.0001	180	0.61	<0.0001	156	-0.46	<0.0001
FEV ₁ /FVC ratio							180	0.89	<0.0001	156	-0.75	<0.0001
MEF25-75 z score										156	-0.78	<0.0001

Table 3.6. Correlations between airway physiology outcomes at study visit when CT carried out. For each main variable (e.g. FEV₁ z score), columns are divided into 3 sub-columns. From left to right for each, these columns represent the number of patients contributing to the analysis, Correlation coefficient (r) from Pearson correlation, and corresponding p value. In view of multiple comparisons, a statistically significant result was considered when p<0.0001.

3.4.2 Structure

As expected, there was a strong correlation between the extent and severity of bronchiectasis (Spearman $r = 0.89$, $p < 0.0001$), (Figure 3.12). To investigate the extent of disease affecting different airway sites (e.g. proximal vs. distal sites), relationships between other CT variables are summarised in Table 3.7. Airway wall thickness (a component of bronchiectasis, and therefore as expected) correlated strongly and similarly with both bronchiectasis scores (extent and severity). Large (and to a lesser extent, small) airway plugging were similarly correlated with bronchiectasis scores. In contrast, air trapping correlated only weakly (correlation coefficient < 0.5) with all other parameters, suggesting either that this may be an independent feature uncoupled from other measures, or that the scoring of this feature is less accurate than others.

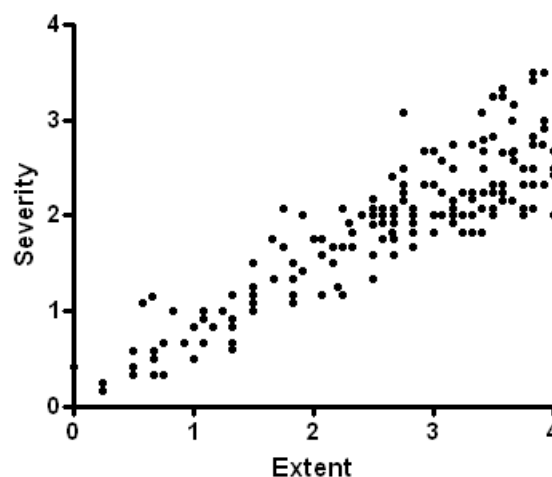


Figure 3.12. Severity and extent of bronchiectasis on CT scan. Each patient is represented by an individual data point, Spearman $r = 0.89$, $p < 0.0001$.

	Severity bronchiectasis			Wall thickness			Air trapping			Small airway plugs			Large airway plugs		
Extent of bronchiectasis	178	0.89	<0.0001	178	0.7	<0.0001	177	0.47	<0.0001	178	0.61	<0.0001	178	0.73	<0.0001
Severity of bronchiectasis				178	0.73	<0.0001	177	0.38	<0.0001	178	0.57	<0.0001	178	0.74	<0.0001
Wall thickness							177	0.37	<0.0001	178	0.54	<0.0001	178	0.7	<0.0001
Air trapping										177	0.35	<0.0001	177	0.36	<0.0001
Small airway plugs													178	0.65	<0.0001

Table 3.7. Correlations between component of CT score and remaining CT outcomes. For each CT variable, columns are divided into 3 sub-columns. From left to right for each, these columns represent the number of patients contributing to the analysis, Correlation coefficient (r) from Pearson or Spearman rank correlation, and corresponding p value. In view of multiple comparisons, a statistically significant result was considered equivalent to $p < 0.0001$.

3.4.3 Inflammation

As the total white count reflected predominantly neutrophils, comparison was restricted to blood neutrophil population and level of serum CRP. For the 175 patients with results for these outcomes at the CT visit (on which all cross-sectional comparisons have been made), there was a weak but significant correlation between neutrophil count and CRP ($r=0.33$, $p < 0.0001$).

3.4.4 Quality of life

Data analysis focused on the CFQ-R Respiratory domain score, and therefore no ‘within category’ comparisons were performed.

3.5 Cross-sectional correlations between categories

To explore relationships, this section compares outcomes between categories (e.g. physiology vs. structure). Again, an investigation of these relationships helps understand issues relating to sensitivity and site of airway or systemic ‘readout’. As before, all measurements included in this analysis are from the visit at which CT was performed.

3.5.1 Physiology vs. structure

FEV₁ z scores, MEF₂₅₋₇₅ and LCI all correlated significantly with all CT parameters; for all physiological measures, the strongest correlation was seen with bronchiectasis and the least strong with gas trapping; mucus plugging was somewhere in between. LCI correlated more strongly with all CT parameters than did any of the spirometric values (Table 3.8). There was no indication from these data that MEF₂₅₋₇₅ was reading out a different level of the airway than was FEV₁. LCI was superior to FEV₁ in correlating with degree of air trapping. The ratio of FEV₁/FVC was almost equivalent to LCI in terms of strength of correlation with air trapping on CT, and was superior to MEF₂₅₋₇₅ (which is conventionally considered to read out small airways).

For all spirometry measures, including MEF_{25-75} large airway plugging was more strongly correlated than small. For LCI, correlations with large and small plugs were much more similar. To investigate the sensitivity of LCI vs. FEV_1 or MEF_{25-75} to detect structural change (as assessed by extent of bronchiectasis), the data was plotted as previously described for comparison within physiology outcomes. Again, LCI was superior in terms of sensitivity to detect change on CT than either FEV_1 z score or MEF_{25-75} (Figure 3.13).

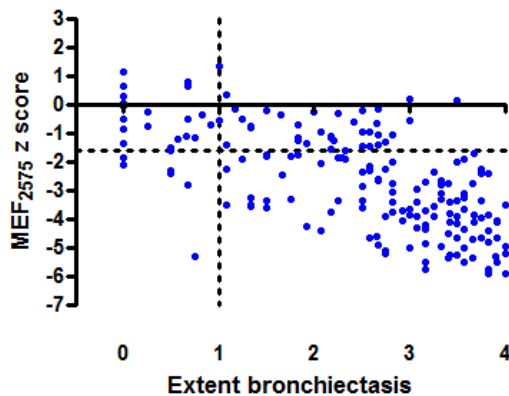


Figure 3.13A

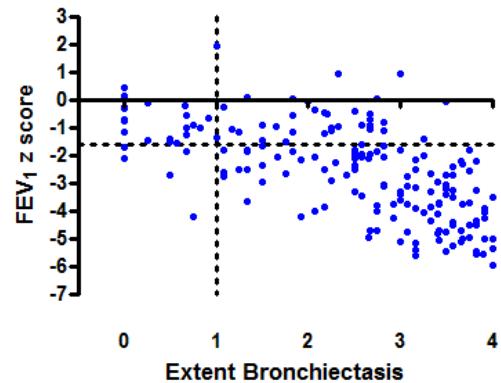


Figure 3.13B

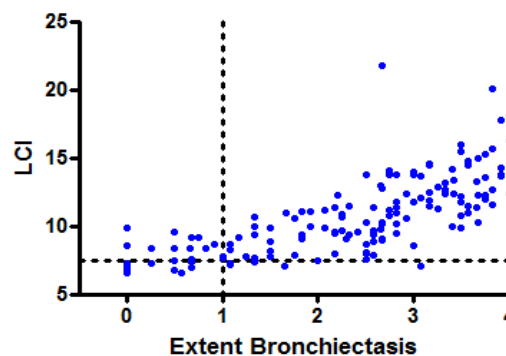


Figure 3.13C

Figure 3.13. Extent of bronchiectasis and FEV_1 , MEF_{2575} and LCI. Figure 3.13A is for extent of bronchiectasis with FEV_1 z score; 3.13B with MEF_{2575} z score and 3.13C with LCI respectively. Each patient is represented by a single data point. Reference lines delineating normal and abnormal results are present at $LCI = 7.5$, FEV_1 or MEF_{2575} z score < -1.64 and a score for extent of bronchiectasis of 1 or more.

	Extent bronchiectasis			Severity bronchiectasis			Wall thickness			Air trapping			Small airway plugs			Large airway plugs		
Physiology																		
FEV ₁ z score	178	-0.68	<0.0001	178	-0.66	<0.0001	178	-0.58	<0.0001	177	-0.42	<0.0001	178	-0.55	<0.0001	178	-0.64	<0.0001
FVC z score	178	-0.54	<0.0001	178	-0.55	<0.0001	178	-0.47	<0.0001	177	-0.31	<0.0001	178	-0.44	<0.0001	178	-0.5	<0.0001
FEV ₁ /FVC ratio	178	-0.61	<0.0001	178	-0.58	<0.0001	178	-0.48	<0.0001	177	-0.48	<0.0001	178	-0.48	<0.0001	178	-0.57	<0.0001
MEF ₂₅₋₇₅ z score	178	-0.67	<0.0001	178	-0.63	<0.0001	178	-0.56	<0.0001	177	-0.44	<0.0001	178	-0.53	<0.0001	178	-0.62	<0.0001
LCI	154	0.82	<0.0001	154	0.78	<0.0001	154	0.65	<0.0001	153	0.5	<0.0001	154	0.64	<0.0001	154	0.68	<0.0001
Inflammation																		
Blood neutrophils	176	0.35	<0.0001	176	0.4	<0.0001	176	0.35	<0.0001	175	0.22	0.004	176	0.33	<0.0001	176	0.39	<0.0001
CRP	174	0.33	<0.0001	174	0.39	<0.0001	174	0.41	<0.0001	173	0.27	0.0003	174	0.22	0.0029	174	0.35	<0.0001
Quality of Life																		
Respiratory domain CFQ-R	178	-0.44	<0.0001	178	-0.42	<0.0001	178	-0.4	<0.0001	177	-0.29	<0.0001	178	-0.35	<0.0001	178	-0.4	<0.0001

Table 3.8. Correlations between physiology, inflammatory and quality of life outcomes with CT outcomes, at study visit when CT carried out. For each CT variable, columns are divided into 3 sub-columns. From left to right for each, these columns represent the number of patients contributing to the analysis, Correlation coefficient (r) from Pearson or Spearman rank correlation, and corresponding p value. In view of multiple comparisons, a statistically significant result was considered equivalent to $p < 0.0001$.

3.5.2 Physiology vs. inflammation

There were weak but statistically significant correlations between FEV₁ and LCI with measures of systemic inflammation (Table 3.9). However significant correlations were not consistently seen across other spirometry outcomes.

3.5.3 Physiology vs. Quality of life

Significant correlations were seen between all spirometric outcomes and LCI with the CFQ-R respiratory domain score (Table 3.9). The strength of correlation between the respiratory score from CFQ-R and spirometry outcomes and LCI was similar to those seen with structural changes. It was greatest for FEV₁, MEF₂₅₋₇₅ and LCI.

3.5.4 Structure vs. inflammation

Blood neutrophil count and CRP measurements were associated with CT outcomes however the relationships were weaker than those described above for airway physiology measurements (Table 3.8). The strongest association between CRP and CT outcome was seen for airway wall thickness. Both blood neutrophil count and CRP were also associated with severity of bronchiectasis on CT scan, and with the degree of air trapping on CT however although remained statistically significant was weaker. A greater association was seen between neutrophil count and severity of bronchiectasis on CT than for wall thickness or air trapping (Table 3.8).

3.5.5 Structure vs. Quality of life

Lower CFQ-R Respiratory symptom domain scores were associated with extent of bronchiectasis on CT (Table 3.8). The relationship between quality of life scores and CT outcomes was greatest for extent of bronchiectasis (in comparison to air trapping and wall thickness).

	FEV ₁ z score			FVC z score			FEV ₁ /FVC			MEF25-75 z score			LCI		
Inflammation															
Blood neutrophils	178	-0.3	<0.0001	178	-0.22	0.003	178	-0.29	<0.0001	178	-0.28	0.0001	154	0.39	<0.0001
CRP	176	-0.34	<0.0001	176	-0.27	0.0002	176	-0.28	0.0002	176	-0.31	<0.0001	152	0.38	<0.0001
Quality of Life															
Respiratory domain CFQ-R	180	0.41	<0.0001	180	0.33	<0.0001	180	0.36	<0.0001	180	0.4	<0.0001	156	-0.4	<0.0001

Table 3.9. Correlations between physiology outcomes and measures of systemic inflammation and quality of life at study visit when CT carried out. For each physiology outcome (e.g. FEV₁ z score), columns are divided into 3 sub-columns. From left to right as before, these columns represent the number of patients contributing to the analysis, Correlation coefficient (r) from Pearson or Spearman rank correlation, and corresponding p value. In view of multiple comparisons, a statistically significant result was considered when p<0.0001

	Blood neutrophils			CRP		
Quality of Life						
Respiratory domain CFQ-R	178	-0.26	0.0006	176	-0.24	0.0015

Table 3.10. Correlations between quality of life and measures of systemic inflammation. The format and statistical significance as defined above.

3.5.6 Inflammation vs. Quality of life

There were no significant correlations between markers of systemic inflammation and the Respiratory domain score from the CFQ-R. Although the results appeared to show a weak association between lower domain scores and higher populations of neutrophils or CRP in blood these did not achieve statistical significance (Table 3.10).

3.6 Longitudinal analysis: change over time

For any medium to long term study of CF, the progressive nature of the disease and thus the potential for directional change independent of any intervention needs to be considered. For the majority of outcome measures in CF, the degree of such directional change has not been defined. This section describes longitudinal data taking into account the number of days between individual study visits. Unless otherwise described, analysis has been performed on patients with results for the outcome of interest at every one of the 4 study visits.

The median number of days between visit 1 and visit 4 for all 150 patients attending 4 study visits was 426 days (interquartile range 399 to 479 days). The minimum and maximum number of days between visit 1 and 4 were 329 and 875 days respectively.

3.6.1 Physiology

3.6.1.1 Spirometric values

The mean (95% CI) change in FEV₁ % predicted per year was -1.1% (-2.2 to +0.02) in absolute % predicted values, which approached statistical significance. Mean change in FEV₁ z score was -0.1 z scores per year. The change in % predicted values per year are plotted in Figure 3.14, with each patient represented by a single data point.

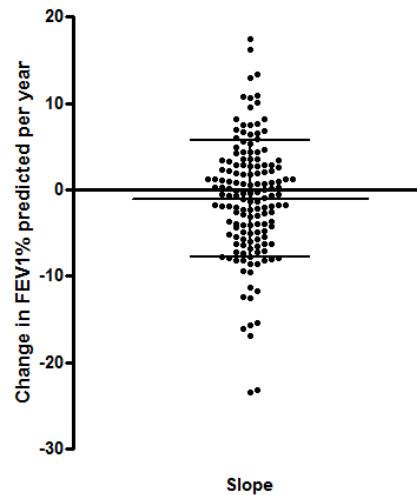


Figure 3.14. Change in FEV₁% predicted per year. Lines represent mean and standard deviation. Each patient is represented by a single data point.

There was no significant change in FVC over time (-0.41% (-1.43 to 0.61) per year) and the corresponding change in FVC z score was -0.04 (-0.14 to 0.05) per year. The mean change in MEF₂₅₋₇₅ % predicted and z scores per year was -1.3% (-3.67 to 1.04) and -0.08 (-0.20 to 0.05) z scores per year respectively, however again this was not statistically significant. There was also no change in the ratio of FEV₁ to FVC over the study period (mean change -0.01% (-0.02 to 0.00) per year). The change in % predicted values per year for FVC and MEF₂₅₋₇₅, and change in ratio of FEV₁/FVC are plotted in Figure 3.15.

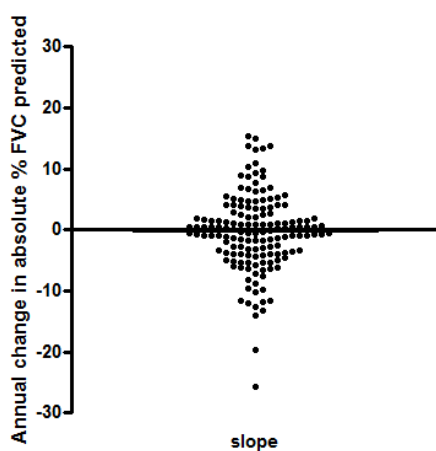


Figure 3.15A FVC % predicted

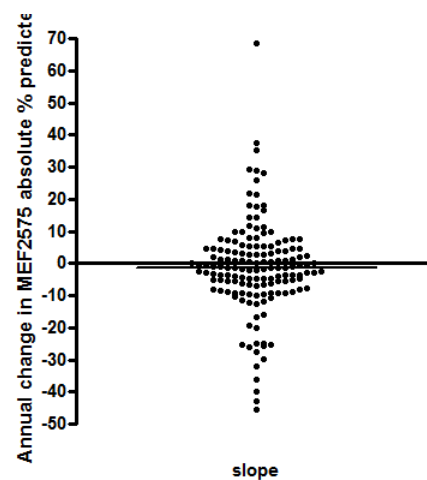


Figure 3.15B MEF₂₅₇₅ % predicted

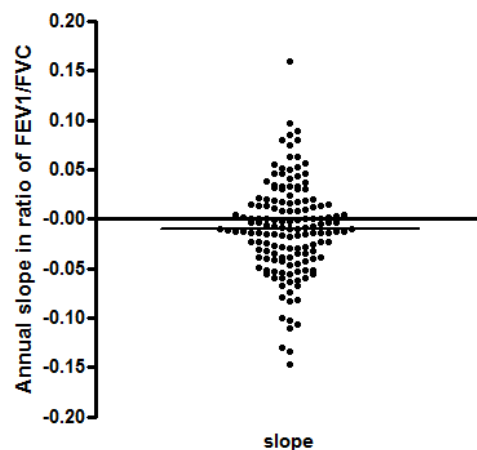


Figure 3.15C Ratio of FEV₁/FVC

Figure 3.15. Annual change in FVC and MEF₂₅₋₇₅ % predicted and ratio of FEV₁/FVC. Presentation of data in graphs A-C as before, lines for mean are largely obscured by the proximity to the x-axis indicating no significant change over time.

3.6.1.2 Lung clearance index

LCI did not change significantly over the study period. The mean (95% CI) change in LCI per year for the n=92 patients with LCI results at each visit was +0.07 (-0.12 to 0.27), summarised in Figure 3.16.

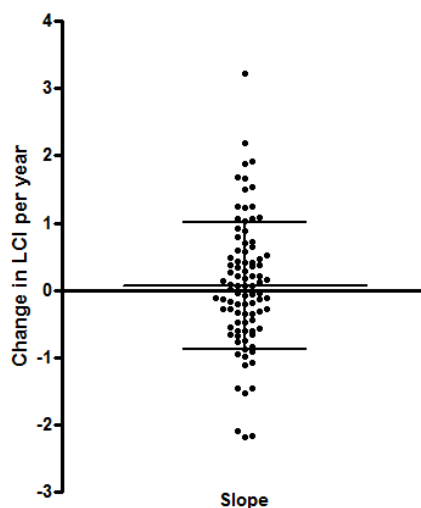


Figure 3.16. Annual change in LCI. Each patient is represented by a single data point as before. Lines represent mean and standard deviation, with no evidence of change over time.

3.6.2 Inflammation

Neither blood total WCC, neutrophils or CRP changed significantly over time. The mean (95% CI) annual change in total WCC, neutrophils and CRP were -0.14 (0.53 to 0.24) $\times 10^9$ /L, -0.04 (-0.44 to 0.35) $\times 10^9$ /L and -0.44 (-1.67 to 0.79) mg/L (Figure 3.17).

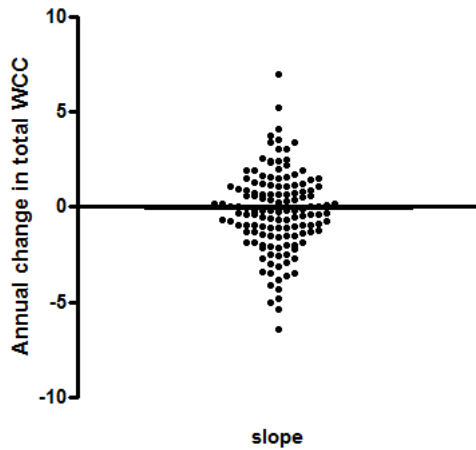


Figure 3.17A Total white cell count

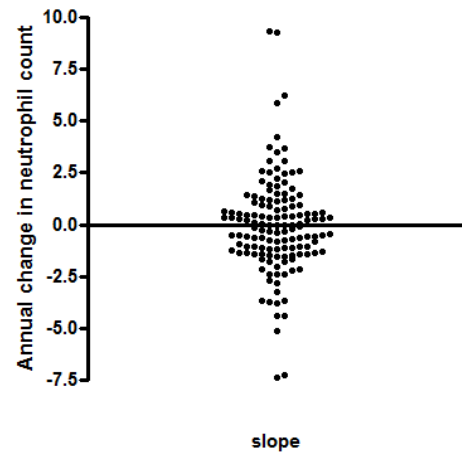


Figure 3.17B Blood neutrophil count

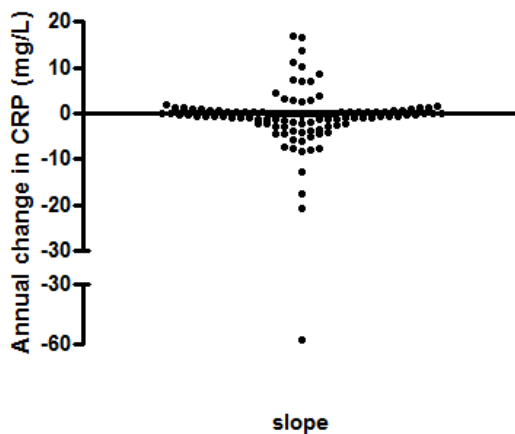


Figure 3.17C C-reactive protein

Figure 3.17. Annual change in blood total white cell count, neutrophils and C-reactive protein. Data is presented as before, however lines for mean change overlie the x-axis at $y=0$.

3.6.3 CFQ-R Respiratory domain

There was no change in patient reported outcomes over time as assessed by the CFQ-R respiratory domain score. The mean (95% CI) change for patients with results from all 4 visits was +0.3 (-1.41 to 1.97) score points per year. Results are displayed in Figure 3.18.

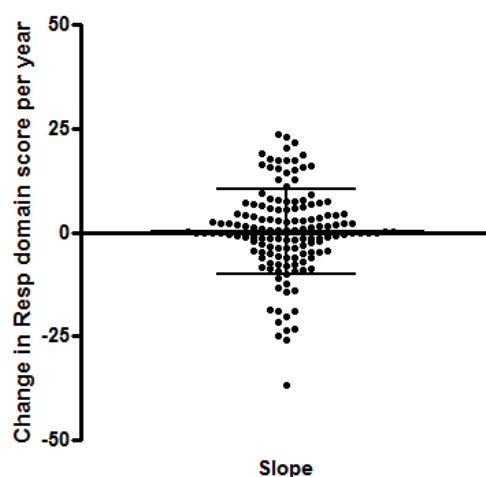


Figure 3.18. Annual change in the Respiratory domain score of the CFQ-R. Data is presented as previous graphs with lines representing mean and standard deviation. Resp = Respiratory.

3.6.4 Influences on longitudinal change

Despite an absence of significant change over time in the outcome measures investigated for the entire study population with data at each of the 4 study visits, the patient population was diverse in terms of disease severity. Change over time, within subgroups was investigated according to visit 1 disease severity (FEV₁ z scores) and age of patient.

The rate of change in FEV₁ was related to the patient's lung function at visit 1. Patients with normal lung function at visit 1 (defined using a FEV₁ z score cut off of -1.64) demonstrated a steeper decline in FEV₁% predicted than those with an abnormal visit 1 FEV₁ z scores (Figure 3.19). Those with normal lung function (n=59) had a mean (95% CI) annual change in FEV₁ of -3.4% (-5.6 to -1.3) in absolute % predicted values compared to a change of +0.22% (-0.9 to 1.4) for those with an abnormal FEV₁ at visit 1 (n=97); unpaired t-test p=0.001. This

relationship was also observed when the Run-In study patient population was divided into two approximately equal groups using a z score cut off of -2.5. Using this divisor, the annual change in FEV₁ was -2.4% (95% CI of mean -4.05 to -0.72) in absolute % predicted values (n=78) for patients in the milder group (z score $\geq$ -2.5) compared to +0.4% (95% CI of mean -0.97 to 1.68) for those in the more severe group (z score < -2.5) (n=72); unpaired t-test p =0.01 (Figure 3.19).

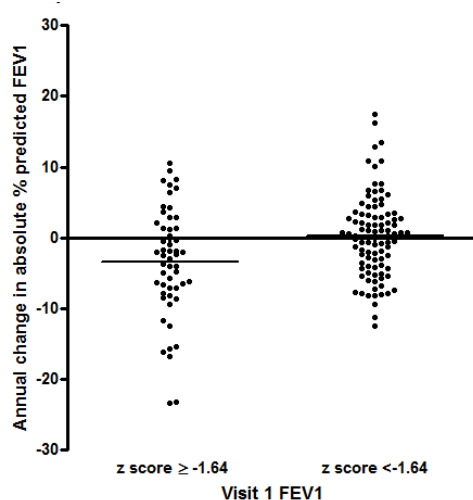


Figure 3.19A

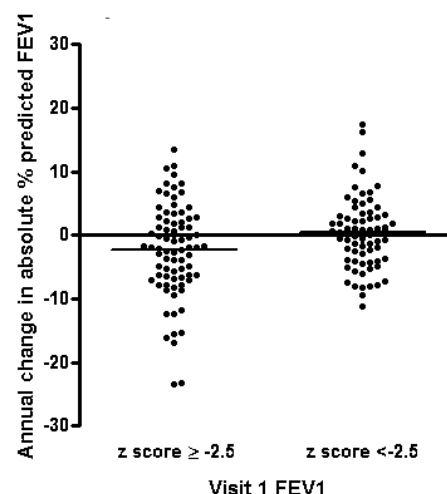


Figure 3.19B

Figure 3.19. Annual change in FEV₁% predicted according to FEV₁ z score at visit 1. The annual change in FEV₁ % predicted is plotted according to a lower limit of normal z score for FEV₁ (-1.64)(Figure 3.19A) and a z score cut off of -2.5 (Figure 3.19B)(which divided the Run-In population into two approximately equal halves). The rate of change in FEV₁ % predicted was greater in both milder groups of patients (unpaired t-test p=0.001 for cut off of -1.64 z scores, and 0.01 for -2.5 z score cut-offs respectively).

There appeared to be a greater rate of decline in FEV₁ z score in patients age < 18 years at visit 1, but this did not reach statistical significance (unpaired t-test for change in z score for patients <18 years vs. $\geq$ 18 years at visit 1, p= 0.13).

However, there was no difference between the trends of LCI or CFQ-R Respiratory domain score over time according to the presence of a normal or abnormal FEV₁ z score at visit 1. The unpaired t-test for mean rate of change in LCI over time according to normal or abnormal FEV₁ z score at visit 1 was not significantly different between groups (p= 0.6), nor was it for trends in change in Respiratory domain score over time (p= 0.3).

3.6.5 Correlation of longitudinal change between outcome categories

If outcome in a clinical trial is assessed according to change in one parameter, this may not be mirrored by 'improvement' in another category. To explore relationships in directional change between one outcome with another over time within individuals, the slopes resulting from the regression analyses were compared.

The annual change in FEV₁ z score was negatively correlated with the annual change in LCI (Pearson $r = -0.34$, $p=0.0009$; $n=92$) (Figure 3.20), and therefore both consistent with a directional change of more severe disease.

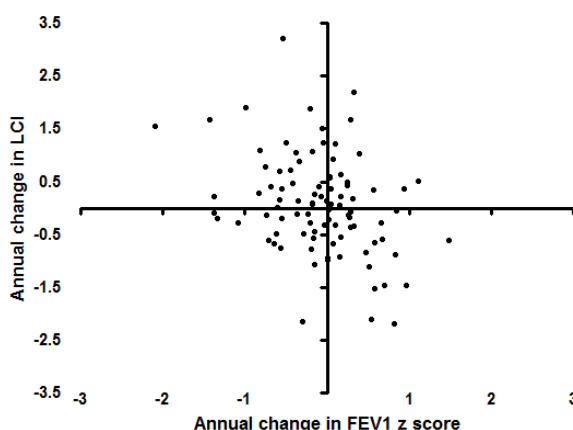


Figure 3.20. Annual change in FEV₁ and LCI. The slope representing annual change in FEV₁ z score against change in LCI per year is plotted for each patient. Pearson $r = -0.34$, $p < 0.01$.

To determine whether rate of change over time in physiological outcomes was related to baseline airway structural abnormalities, comparisons were made with baseline CT parameters. There was no significant correlation between the annual change in FEV₁ z score or LCI and the severity of bronchiectasis score on chest CT (Spearman $r = 0.1$, $p = 0.2$ and $r = -0.05$, $p = 0.6$ respectively). Similarly there was no association between trends in FEV₁ z scores and extent of air trapping or bronchial wall thickness on CT. Although direction and magnitude of change in physiological outcomes correlated with each other, they also did not reflect any change in the CFQ-R respiratory domain score over time. There was no correlation between the direction and rate of change in FEV₁ z score or LCI

and direction of change in the Respiratory Symptom domain score of CFQ-R (Pearson $r = -0.07$, $p=0.4$ and $r=-0.17$, $p=0.1$ respectively), (Figure 3.21). Any within-patient trends over time in spirometry or LCI were not reflected by similar trends in inflammatory markers (CRP, white cell count and neutrophil count) (data not shown).

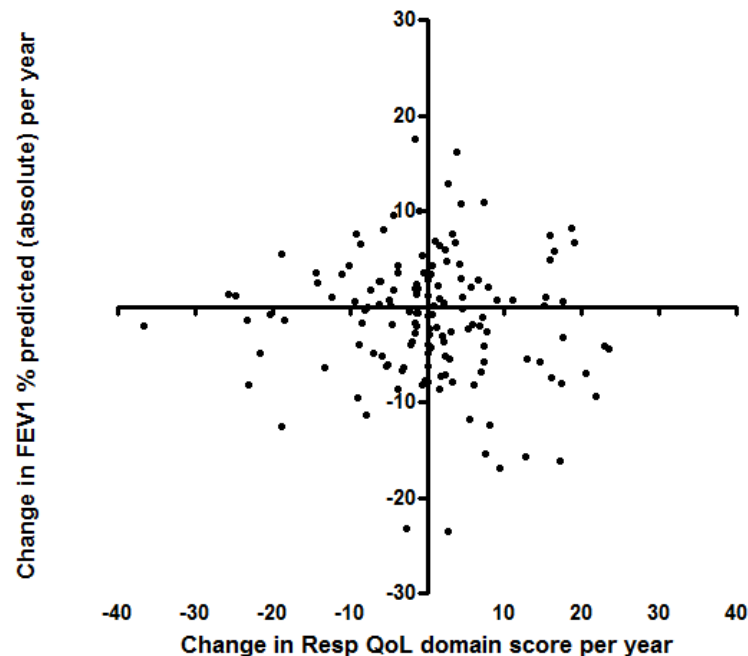


Figure 3.21. Annual change in Respiratory domain score of CFQ-R and FEV₁ z score. Format of graph is as presented for other comparisons. No significant correlations were observed (Pearson $r=-0.07$, $p=0.4$).

3.6.6 Change over time with data from < 4 study visits

Inclusion of data from patients with 2 or more study visits (as opposed to 4) increased the ‘n’ number of subjects per outcome measure but also increased the variability of change over time when included in the regression analyses. Therefore results with <4 data points have not been described within this section. An example of the potential problems with a small ($n=2$) number of data points to plot a ‘best fit’ line to represent change over time is illustrated by the CFQ-R respiratory domain score results. The change per year for all patients with 2 or more visit results according to regression analyses is shown in Figure 3.22. This suggests an annual change beyond the potential maximal value possible for this outcome (for example by having data from 2 visits only and a

CFQ-R respiratory domain score of e.g. 33 at visit 1 and 68 at visits 2 separated by a 3 month interval) and is therefore invalid.

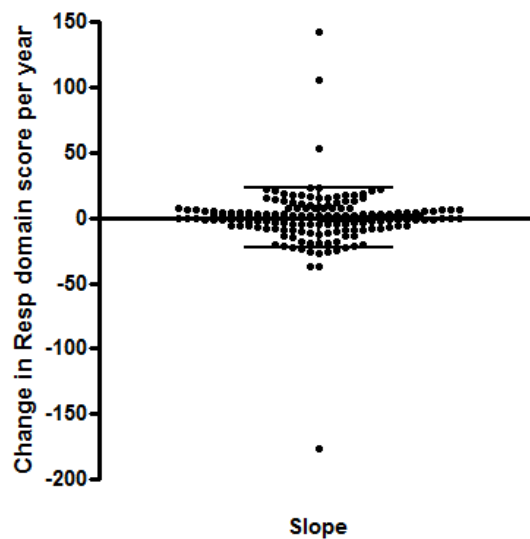


Figure 3.22. Annual change in Respiratory domain score of CFQ-R calculated according to slope derived from 2 or more (rather than 4) study visits. The resultant ‘change’ in score over time for the outliers is implausible (maximum domain score =100). Respiratory is abbreviated to ‘Resp’; Lines represent mean and standard deviation as previously.

3.7 Discussion

The outcome measures described for the Run-In study have included data analysis from two perspectives: 1) results at each visit in order to describe variability in different outcome measures over time, and 2) results according to the actual visit interval in order to investigate rate of change in outcome over time (sections 3.3 and 3.6 respectively). Interest in the former is important to guide suitability as outcome measures for clinical trials and to guide power calculations. The second allows an investigation of directions change over time indicating progression of lung disease. Results from cross-sectional comparisons both within and between categories of outcome measures have also been described (section 3.4 and 3.5), allowing investigation of their relationship with each other.

Several themes are worthy of further discussion. These are related both to general issues of longitudinal study design and specific issues to the outcome

measures described, including methodology as described in Chapter 2, section 2.3-2.4.

3.7.1 Airway physiology analysis methodology

The results for the Run-In study airway physiology outcomes in this chapter are a consequence of refinements initiated as part of the UKCFGTC clinical programme as described in Chapter 2, section 2.3. Through involvement in both data collection and data analysis, the author was able to identify several issues within physiology outcomes that were being collated and speculate on their likely cause. Collectively, the processes employed have provided opportunity to investigate the impact of variables on final outcomes which contribute to data analysis for clinical trials. These include reference equations, test software and hardware, and other external factors which can influence final results.

Unexpected results, including discrepancies on repeated analyses of the same test (as identified for LCI) or exaggerated longitudinal trends over time (as identified for FEV₁ % values) initiated the process of review. As a result of the exhaustive team effort between physiologists and research fellows at the London and Edinburgh sites in LCI reanalysis, we were collectively able to identify additional problems with quality of test data which led to refinement in methodology, and detailed revision of standard operating procedures. The cost was data loss - most significantly at the London site - but the resultant data surviving quality control measures can be analysed with greater confidence as a consequence. Horsley reported that > 85% of CF subjects performing an Innocor-measured LCI were able to complete testing and obtain reproducible measurements (48). The 97% rate for LCI test acceptability at the Edinburgh sites were therefore impressive and, in combination with a standardised approach and environmental improvements in London testing conditions, is encouraging for superior dataset completeness in the ongoing multidose trial.

Moving forward, it will be important to determine whether the finding that a greater proportion of LCI tests are rejected in patients with more severe disease holds true. If so, this is crucial information for planning power calculations for any CF clinical trial including LCI as an outcome measure in patients with moderate or severe lung disease. Within-patient variability in LCI during longitudinal data collection is described in section 3.3.2; higher coefficients of

variation were observed for the most severe quartile of Run-In patients. Collectively these issues suggest that as an outcome measure LCI is better suited to those with early stage disease, or at least consideration should be given to determining an appropriate treatment effect size in interventional trials according to the underlying severity of disease.

During the process of data analysis for the Run-In study, the author became aware that the use of the two separate spirometry reference sources (Rosenthal and Quanjer) was introducing unacceptable additional variability into longitudinal assessment of change in % predicted values over time (Chapter 2, section 2.3). Presentation of the comparative data detailed below to the UKCFGTC Strategy Group led to the adoption of Stanojevic 'All ages' approach for interpretation of spirometry values for both the Run-In and Pilot studies, and this continues to be used in the current multidose trial. The conversion of spirometry data to % predicted values highlighted the challenges introduced by an external reference source, particularly with separate, non-continuous reference ranges for children and adults. For studies in a disease such as CF which may include participants in their teenage years, use of discontinuous reference ranges across child- and adulthood can have significant implications for data interpretation and clinical decision making. For longitudinal studies including this patient group a continuous reference source is essential. The opportunity to take account of expected distributions within a population by defining results according to z scores is also attractive, as the implications of a particular % predicted value are not static with age. In the future the author anticipates similar 'all age' reference sources being crucial to appropriate LCI interpretation as exists for spirometry (64) currently.

3.7.2 Recruitment and retention

Interestingly, the rate of patient withdrawal between Run-In study visits was consistent, demonstrated by the almost perfect regression line across data points representing patient numbers at each visit (section 3.1.2). The interval between study visits was not constant, increasing with later Run-In visits (Figure 3.2). This reflects the difference in perceived vs. real timelines that were encountered in preparation for the UKCFGTC Pilot study (single dose safety trial, see Chapter 2, section 2.2) and multidose gene therapy trial and a resultant

spacing of Run-In visits. Maintaining patient enthusiasm for on-going participation in longitudinal studies is a challenge. Although its wider applicability is unknown, the linear relationship between visit number and patient retention is a useful observation for non-interventional studies leading to patient selection for a future (interventional) trial such as intended by Run-In study and its relationship with the multidose trial.

3.7.3 Defining clinical stability

In the Run-In study, a definition of clinical stability was based on patient history and a crude (in-house designed) symptom score. Patients were only allowed to continue with the study visit if the symptom score was compatible with no or little change from baseline clinical status. Despite this, it is interesting that individual patients can be identified as significant outliers for certain tests. For example, the serum C-reactive protein (CRP) results demonstrate a relative stability over time, however on a single occasion an individual patient had a CRP above 100mg/L. Other examples of significant outliers observed within the results included scores for the Respiratory symptom domain of the CFQ-R from a single visit being well outside those from the other 3 visits analysed.

Therefore it is possible that other outcomes measured at the same visit may not represent that patient's 'true' baseline status, and may artificially affect the expected variability of a test. Ultimately, a pragmatic approach has to be taken and in accordance to data presented and discussed in this chapter, an inclusive approach has been undertaken. All data were analysed, irrespective of the presence of a significant outlier.

3.7.4 Data availability

The author has chosen to concentrate longitudinal analysis on those patients with a complete data set (results from 4 visits) for the outcome of interest. This inevitably meant that not all patients were represented in the analysis. This was due either to patient withdrawal from the study prior to visit 4, or to absent data as a result of missing test results or exclusion of results prior to inclusion on the CTEU database as a result of quality control processes (e.g. as occurred for LCI, described in Chapter 2, section 2.3). Comparison of the visit 1 results from

these patients with the entire study population at visit 1 demonstrated that they were not significantly different at baseline, with the exception of LCI. This was significantly higher in the overall study population. Although there were no differences in other outcomes, it is possible that concentrating on complete datasets excluded patients with more severe disease and this is an important consideration when considering change over time, however the additional variability introduced by including patients with incomplete datasets meant that this was an accepted limitation. However, it may be that the difference is artefact, as a higher proportion of LCIs from patients with more severe disease were ineligible for the final Run-In LCI dataset (Chapter 2, section 2.3).

3.7.5 Coefficient of variation

The measure ‘coefficient of variation’ (CV, which is calculated by dividing the standard deviation by the mean) has been calculated for the longitudinally collected outcome measures reported in this chapter (section 3.3). The median coefficient of variation between visits was 6.0% for FEV₁% predicted values and 6.4% for LCI. Tests with low expected values, irrespective of the units, are subject to large coefficients of variation. Therefore results which are maintained in a ‘normal’ range during periods of stability can still experience wide variability (and high coefficients of variation) between visits. This was illustrated by the results across visits 1-4 for white cell count and CRP. Coefficients of variation, as a reflection of variability of an outcome over time, are also not suitable for assessing variability in z scores between visits as they contain both positive and negative values. Another difficulty with this summary measure is that an outcome with a small unit of measurement and an apparently ‘low’ CV (%) can result in an expected variability in the outcome which is equivalent or near to the magnitude of treatment effects seen in response to a therapeutic intervention. For example, a CV of 7.0% on the basis of an LCI of 10 is equivalent to an absolute value of 0.7. This is near to the size of treatment effect with RhDNase in the study reported by Amin et al (70) (LCI was measured in 17 patients in a cross over trial involving RhDNase or placebo; mean (SD) improvement in LCI on RhDNase = 0.9 (1.44)). CVs do however provide useful information about the variability of outcomes over time within the patient population and can help guide power calculations for determining appropriate sample sizes in interventional clinical trials.

3.7.6 Ceiling effect

The results from the CFQ-R demonstrate the problem of ceiling effects with outcome measures. The CFQ-R domains for physical functioning, role, eating disturbances and weight were particularly affected by this with over half of all study participants attaining a score >90 (maximal score 100), as shown in Figure 3.9. Furthermore, for 7/12 domains in the adult questionnaire, the 75th centile was >90. If the patient population of interest has a maximal or near maximal score when well during periods of stability, there is little opportunity to observe a treatment effect. This is even more the case with CFQ-R results than lung function measurements, as the latter are based on a continuous numerical scale with no maximum value in comparison to an absolute maximum of 100 for each CFQ-R domain. These ceiling effects also reduce the validity of plotting the domain score against a continuous variable such as FEV₁% predicted, as linear associations become invalid if the scores are saturated at the top of the scale. As scores from the Respiratory domain were not all saturated at the top of the scale, linear correlations were made in this thesis. Other potentially useful domains in the CFQ-R to explore further in this dataset include Vitality, Social, Health perception, and Treatment burden. An association between these domains and disease severity in CF using the CFQ has been reported in the literature (106).

3.7.7 Change in outcome measures over time

Data analysis of repeated measures from individual subjects was undertaken in order to identify change over time (section 3.6). Whereas the assessment of variability was undertaken independent of visit interval (as Run-In visits were undertaken at times of clinical stability defined by symptom score undertaken at study visit), trends in change over time were undertaken by taking the visit interval into account.

The assessment of change over time was performed with linear regression for each outcome independently. However, relationships presented here suggest there would be merit in performing sophisticated statistical modelling analysis to identify the interdependence of the different outcomes investigated throughout this chapter. Such models could also take into account patient characteristics

including airway microbiology, concurrent medications, and infective exacerbation rate requiring admission for intravenous antibiotics over the data collection period.

Despite some wide variations in 'slopes' from the regression equation (representing change over time) from particular outcome measures between individual patients, for the study population there was little change in the results from airway physiology tests, blood markers of infection and inflammation, and quality of life over the duration of the first 4 Run-In study visits. The mean decline in FEV₁% predicted values per year in the Run-In population was consistent with reports in annual decline in FEV₁ for patients with CF described in the literature (46). Steeper rates of decline in FEV₁ in the Run In study were observed in patients with milder disease (as assessed by FEV₁) at visit 1. This is in keeping with the findings of Konstan et al where high baseline FEV₁ was an independent risk factor for decline in FEV₁ % predicted in children and adolescents (118). The comparatively lower rate of decline in FEV₁ over time in patients with more severe disease is also an important consideration for CF clinical trials when the primary outcome is decline in z scores or % predicted values for FEV₁ over time. The use of a continuous reference source for paediatric and adult patients (as described in Chapter 2, section 2.3) has allowed a more appropriate investigation of change over time than if two discontinuous reference sources were used - longitudinal change in % predicted FEV₁ must not be assessed with the latter. Use of separate paediatric and adult reference ranges for % predicted values from the Danish CF Registry have recently been used to model the rate of decline in FEV₁ % predicted values according to baseline FEV₁ % (119). I suggest that any prediction model developed from this discontinuous dataset should be interpreted with caution.

Of interest was the finding that the direction of change (in terms of slope of regression line) for FEV₁ did not correlate with the annual change in Respiratory symptom domain score for CFQ-R on an individual patient basis. This provides evidence for the importance of measuring multiple outcomes in clinical studies involving CF patients. It is possible that the change over time would better conform to expected direction of change with longer periods of follow up. Data collection over a more prolonged period of time (spanning multiple years, if not decades) is needed to evaluate a more realistic description of annual rate of

change than possible with the data analysed here. However, from a clinical trial perspective this is impractical.

3.7.8 Relationship between structural change on CT and other outcomes

Even when corrected for multiple comparisons, statistically significant correlations were seen between measures of disease severity as assessed with airway physiology tests (spirometry and LCI) and structural assessment of the airway with CT including assessment of bronchiectasis, airway wall thickness and degree of air trapping. However, the strength of these associations was not uniform. Due to the multitude of 'significant' p values resulting from correlation of CT outcomes with spirometry, the reader is directed instead to the strength of association suggested by the correlation coefficients. Consistent with the literature, LCI was superior to FEV₁ and other spirometry indices in the correlations with structural change as assessed on CT (120). It was more sensitive than either FEV₁ or MEF₂₅₋₇₅, and in support of the study by Gustafsson et al we also found that a normal LCI almost excludes the presence of bronchiectasis on CT (120). The author did not perform a sensitivity analysis on the degree of air trapping with physiology outcomes as all CT scans performed were scored as having at least 30% air trapping. This is the % cut-off often used for defining abnormality (121) and was used as the threshold for abnormal gas trapping in the sensitivity analyses between airway physiology and CT outcomes performed by Gustafsson (120).

The strength of correlation seen between FEV₁ and structural assessments on CT were similar to MEF₂₅₋₇₅ for all CT outcomes. Therefore there was no indication from these data that MEF₂₅₋₇₅ was reading out a different level of the airway than was FEV₁. For all spirometry measures, including MEF₂₅₋₇₅ (which is thought to read out small airways), large airway plugging was more strongly correlated than small. For LCI, correlations with large and small plugs were much more similar. The highly statistically significant correlations between LCI and severity of bronchiectasis on CT support its role in detecting structural lung disease in CF.

3.7.9 Patient reported outcomes

In the Run-In study, there were significant correlations between the Respiratory domain of CFQ-R and all CT outcomes, and similarly with all airway physiology outcomes. In comparison there was no correlation between the change over time in Respiratory domain score from CFQ-R and airway physiology outcomes. This was perhaps not unexpected as one is a patient reported outcome and the other a surrogate marker of the underlying pathology within the lung. This highlights the importance of including patient reported outcomes and supports the use of both in clinical trial design (2).

3.7.10 Variability according to disease severity

In CF the 'optimal' outcome measure is different for different stages of disease. For example, LCI was superior to FEV₁ in relation to the detection of disease on CT, however of interest was the suggestion of greater variability with increasing severity of disease. The increase in variability of the CFQ-R Respiratory domain scores were related to disease severity as defined by FEV₁ z scores, again an important finding for clinical trials. These findings are important for any power calculations aiming to determine sample size for a particular effect change in the outcome measure of interest.

3.7.11 Ability to detect change

The airway is not a single entity and therefore it is important that clinical trials include tests that can detect changes at different airway sites. The relationship between patient reported outcomes and airway physiology tests in longitudinal study is likely to be complex. Scores for the Respiratory domain of CFQ-R have been reported in the literature to be associated with disease severity (106) (and significant correlations with CT parameters, spirometry outcomes and LCI have been reported here for the Run-In study), however the absence of any relationship between direction of change over time for quality of life and spirometry or LCI outcomes was striking. The duration of the Run-In study was not sufficient to investigate whether there is a measurable time lag between physiology and patient reported outcomes.

3.7.12 *Moving forward to multidose trial*

A review of the Run-In data by the trial statistician and UKCFGTC strategy group was undertaken in order to identify the most appropriate outcome measures (and most appropriate patients) to take forward to multidose trial. This led to the selection of relative change in % predicted FEV₁ after 12 monthly doses of CF gene therapy or placebo as the primary outcome for the multidose trial. However, all outcome measures described in this chapter are being collected at intervals throughout the multidose trial. In combination with results from the Pilot study (Chapter 4), this will provide extensive further opportunity to investigate the associations observed between outcomes in the Run-In study and how they change in relation to one another as a result of a novel therapeutic intervention.

4 Results: Single dose Pilot study

This chapter describes patient demographics, baseline clinical status and safety outcomes from the Pilot study. All results from measurements of airway potential difference (PD) in the Pilot study from the upper and lower airway are described in Chapters 5 and 6 respectively.

4.1 Recruitment

Patient enrolment in the Pilot study commenced in February 2009. The final patient was enrolled in May 2011. Forty two patients were enrolled in the study and 36 received a dose of CF gene therapy and completed the trial. Six patients withdrew from the study prior to receiving the gene therapy dose or did not meet clinical stability required for dosing during the time course of the trial.

4.2 Baseline characteristics

Thirty five patients received a single lung dose of gene therapy and n=22 received a nasal dose. With the exception of one patient, all those receiving a nasal dose were also in the lung cohort. This patient did not meet the inclusion criteria for the lung cohort, hence participation in the nasal cohort alone.

4.2.1 Lung gene therapy patients

The median age of all patients receiving a lung dose of gene therapy (n=35) was 27.4 years (range 16.4 - 61.6 years), 34.3% were female. At enrolment, all patients had an FEV₁ % predicted >60% according to Quanjer or Rosenthal (<18 years) reference values (61;62). The median (IQR) FEV₁ % predicted values at enrolment was 72.1 (66.2-86.8) % according to the Stanojevic reference source (64).

Table 4.1 summarises the baseline characteristics of the patients receiving a nebulised dose of gene therapy.

	Yes N (%)	No N (%)
Pancreatic insufficient?	29 (82.9)	6 (17.1)
Nebulised antibiotics?	28 (80)	7 (20)
RhDNase	14 (40)	21 (60)
Hypertonic saline	8 (22.9)	27 (77.1)
Daily (long term) azithromycin	23 (65.7)	12 (34.3)
CF related diabetes on insulin	6 (17.1)	29 (82.9)
Regular inhaled steroids	16 (45.7)	19 (54.3)

Table 4.1. Baseline characteristics of the n=35 patients receiving a nebulised dose of CF gene therapy.

The genotype of patients in the lung cohort was as follows:

- 20/35 (57.1%) were homozygous for the p.Phe508del mutation.
- 14/35 (40.0%) had one known p.Phe508del mutation, with:
 - N=7 another CF disease causing mutation
 - N=7 an unknown mutation.
- One patient had two other (non p.Phe508del) disease-causing mutations.

4.2.2 Nasal gene therapy patients

The median age of the n=22 patients receiving a nasal dose of gene therapy was 26.8 years (range 16.4 - 61.6 years), 37.5% were female. Sixteen patients in the nasal dose cohort were dosed during part A and 6 in part B of the Pilot study.

All patients in the nasal cohort had at least one p.Phe508del mutation. The genotype of patients in the nasal cohort was as follows:

- 12/22 (54.5%) were homozygous for the p.Phe508del mutation.
- 10/22 (45.5%) had one known p.Phe508del mutation, with:
 - N=5 another CF disease causing mutation

- N=5 an unknown mutation.

4.3 Dose delivery and duration

The time taken to deliver each 5ml aliquot of gene therapy to the lung via nebulisation was consistently around 25 minutes when delivered at a ‘standard’ rate. The proportion of complex administered via nebulisation (lung dose) for each dose cohort is shown in Figure 4.1. The mean (SD) volume administered for patients in the 20ml cohort was 17.8 (0.35) ml. Similarly for the 10ml cohort mean volume was 8.74 (0.23) ml and for the 5ml dosing cohort 4.47 (0.23) ml. The proportion of intended dose administered was consistent across the 20ml, 10ml and 5ml dose cohorts (mean 87-89% for each group; ANOVA $p=0.21$).

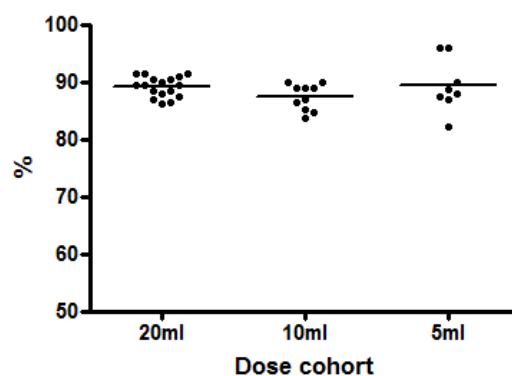


Figure 4.1. Proportion of nebulised gene therapy complex administered according to dose cohort. The proportion of dose administered was determined according to change in weight of nebuliser pot pre- and post administration; lines represent mean values, ANOVA $p=0.21$.

For patients prescribed a 1ml nasal dose ($n=3$), container weights pre and post dosing were available for 2 patients, allowing determination of delivered volume. They received 1ml and 0.8ml respectively. No weights were available for the first patient dosed as the decision to weigh nasal containers in addition to the nebulisers (which had been planned *a priori* to determine actual nebulised volume administered) was made following this. Weights were available to calculate the nasal volume administered for 18/19 patients prescribed a 2ml nasal dose. The mean (SD) volume administered was 1.89 (0.3) ml. The proportion of complex administered for each dose cohort is shown in Figure 4.2.

The variability in actual volume administered was large (as it was based on number of (patient) administered sprays rather than absolute change in weight). For one patient, despite delivering the 24 nasal sprays (as per protocol for 2ml dose), the actual volume administered was only 0.97ml. The maximum nasal dose received was 2.3ml. Sprays were administered to both nostrils in equal proportion.

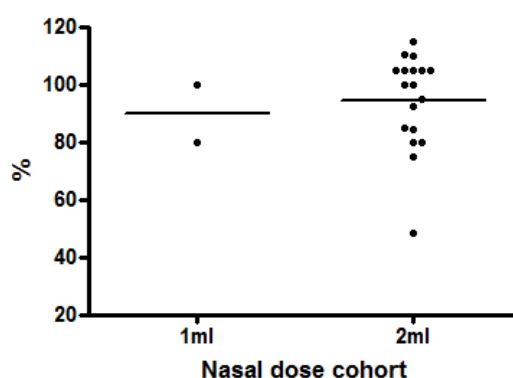


Figure 4.2. Proportion of nasal gene therapy dose administered according to dose cohort. The proportion of dose administered was determined according to change in weight of nasal spray bottle pre- and post administration; lines represent mean values.

4.4 Post dose follow up

4.4.1 Pilot A: 10ml lung dose/1ml nasal dose patients (n=3)

For all three patients, lung and nasal dosing was well tolerated acutely, with report of throat discomfort only. Spirometry immediately post dosing demonstrated a mean drop in FEV₁ of 15% in comparison to immediate pre-dosing values, but this was accompanied by minimal chest symptoms. There was a further decline in FEV₁ and FVC noted within the 6 hours of post-dosing observations, accompanied by findings of crepitations on chest auscultation but no wheeze. There were small drops in oxygen saturation but no patient required supplementary oxygen. Systemic symptoms included low-grade fever, headache, muscle ache and tiredness. A raised white cell count with a predominant neutrophilia was detectable by 8 hours post dosing in all patients. All 3 patients were discharged home the following day (day 1) after elective overnight

observation. The maximal rise in the acute phase protein CRP (as a measure of systemic inflammation) was detected at day 2. Summary graphs of clinical, airway physiology and systemic inflammatory outcomes are presented later in section 4.5, with all patients grouped according to volume of lung dose received.

4.4.2 Pilot A: 20ml lung dose/ 2ml nasal non-invasive lung dose patients (n=3)

Due to the post-dosing profile experienced by the three 10 ml patients, an amendment to the trial protocol (version 5, 06 April 09) was submitted in order that the first three 20 ml patients would not undergo a post dosing bronchoscopy, which could interfere with our understanding of the host response to the agent.

The spectrum of symptoms and decline in lung function was as observed for the 10 ml patients, but was of a greater magnitude. Decline in respiratory status and systemic inflammatory response was more marked (maximum drop in FEV₁ around 40%, maximum temperature post dosing 39.5 degrees C). Again, the clinical picture was of crepitations rather than wheeze and there was little response to bronchodilators (as may have been expected as there was a concomitant drop in FVC). In comparison to the objective measures of lung function drop, the report of respiratory-related symptoms by the patients was less notable. All three were discharged home the following day (day 1).

A day 2 post dosing CT scan on one of these patients detected small areas of 'ground glass', not present on pre-dosing scans. This was considered likely to represent a small patch of either alveolitis or very distal bronchiolitis. In view of this, and the unknown natural history of its resolution, a decision was made by the Data safety monitoring board (DSMB) to withhold further gene therapy dosing until these 3 x 20ml lung dose patients had completed the study protocol. We noted a longer 'tail' of symptoms in the 20 ml patients, with symptoms of tiredness persisting longer than seen in those given 10 ml. The CT changes seen at day 2 in the one isolated patient had resolved by day 14, and inflammatory markers had also normalized by this time point (peak CRP at day 2).

4.4.3 Pilot A: 20ml lung/ 2ml nasal invasive cohort (n=10)

Following discussion of the safety profile of the 3 initial 20ml patients, permission from the DSMB was granted to proceed to the invasive (bronchoscopic) arm of the study. Due to the now familiar post dosing inflammatory profile seen, a further amendment was requested (protocol version 6, 15/07/09) from the regulatory authorities to change the timing of the post dosing bronchoscopy from day 2 and day 14 (originally planned for n=12 in each group) to day 6 and day 14. This was changed to prevent patients experiencing a likely post-gene therapy inflammatory process from being further compromised by a bronchoscopy under general anaesthetic at the same time point.

Ten patients were dosed in this arm of the study. A consistent pattern of clinical and laboratory responses as observed in the 20 ml non-bronchoscopic patients was seen, although the magnitude of change in clinical and laboratory markers between patients was variable. The clinical response and logistics (e.g. distance lived away from hospital) dictated whether individual patients were discharged home or remained in hospital until the following morning. Treatment with ibuprofen and/or paracetamol was given in response to systemic symptoms or fever (see section 4.7). No further post dosing changes were seen on any CT scans. In addition to the raised inflammatory markers, there were small rises detected in serum creatinine and unconjugated bilirubin in the 24 hours post dosing, which then returned to pre-dosing levels. The change in serum creatinine and unconjugated bilirubin remained within the normal range across all measurements for the vast majority of patients despite this small rise from pre-dosing levels, and are considered likely to reflect clearance mechanisms of the gene therapy product. Due to the adverse events noted in the 20ml group post dosing, the 20ml invasive bronchoscopic cohort was limited to 10 patients.

4.4.4 Pilot B: dose cohorts

As described, the aim of part B of the study was to find an optimal dose to take forward to multidose trial. Lung dosing strategies included:

- 10ml lung dose (slow delivery), 20ml lung dose (slow delivery)

- 20ml lung dose with adjuvant prednisolone or ibuprofen
- Further standard rate dosing with 10ml and 5ml volumes
- 10ml or 5ml lung doses with adjuvant paracetamol

Patients in Pilot B experienced the same pattern of symptoms and changes in laboratory markers post dosing as those described for part A although the magnitude of this response was related to the dose received for most markers. In part B, n= 3 patients were given their lung dose at a slow delivery rate; n=2 receiving a 20ml lung dose and n=1 receiving a 10ml lung dose. Slowing the rate of delivery made no difference to clinical or laboratory outcome parameters and the results from these subjects have therefore been combined with patients receiving standard rate nebulisation. All 5ml patients received the dose delivered at standard rate.

4.5 Summary post-dose results according to cohort

To allow comparison between dose cohorts (according to nebulised volume) and the influence of adjuvant therapy, result from patients from part A and B were combined according to lung dose irrespective of the rate of delivery. A summary of the safety outcomes for all patients without planned adjuvant therapy (which are discussed separately below) are summarised here.

4.5.1 Clinical

The maximum temperature was plotted at 8 hours post dose according to dose cohort (Figure 4.3). This was performed as not all patients stayed overnight post dosing and therefore there was an incomplete data set for day 1 results; additionally, for most patients, this was the period of maximal fever.

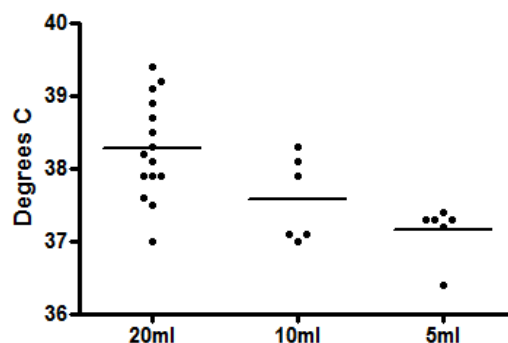


Figure 4.3. Temperature 8 hours post start of lung dosing according to dose cohort. Each patient is represented by an individual data point; patients with planned adjuvant therapies are not included. The temperature post dosing was significantly different between dosing groups (Kruskal-Wallis test $p=0.006$).

4.5.2 Physiology

Spirometry demonstrated a dose dependent difference in maximal decline in FEV_1 post nebulised gene therapy (Figure 4.4), which was detectable 4-6 hours post start of dosing. The subjective report of symptoms by patients did not mirror the magnitude of drop in lung function, even for 20ml patients with a 40% drop in FEV_1 volumes.

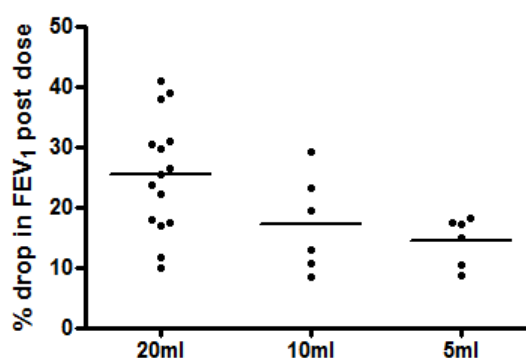


Figure 4.4. Maximum drop in FEV_1 (as % of raw volume) on dosing day. There was a significant difference between dosing cohorts (Kruskal-Wallis test $p=0.03$).

The ratio of FEV_1/FVC remained largely unchanged, suggesting that the underlying processes were causing a restrictive rather than obstructive lung defect. A significant decline from baseline in gas transfer was present on day 2 in patients receiving a 20ml lung dose (paired t-test $p<0.01$)(Figure 4.5). This

had resolved by day 14; there were no significant differences between baseline measurements and those at day 14 or 28 (paired t-test $p=0.26$ and 0.69 respectively). No significant changes were seen between baseline and post dosing KCOc results for the 10ml or 5ml dose cohorts (paired t-test $p>0.05$ for all tests).

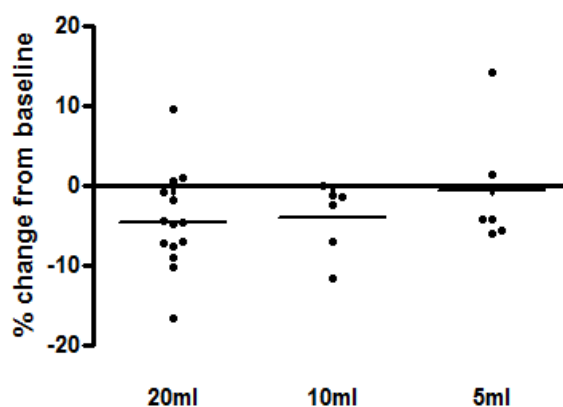


Figure 4.5. Change in KCOc between baseline and day 2 according to dose cohort. There was a significant reduction in KCOc between baseline and day 2 test for the 20ml group (paired t-test $p<0.01$) but not for the 10ml or 5ml dose cohorts.

The LCI signal from the 20ml dose cohort was suggestive of an adverse response. In contrast, patients in the 5 and 10ml dose cohorts showed a reduction in LCI over the study period (mean pre-LCI in 5ml and 10ml dose cohort combined vs. day 28 LCI $p=0.039$)(Figure 4.6). The paired t-test for mean pre- LCI in 5ml cohort vs. day 14 LCI was $p=0.006$ (Figure 4.6).

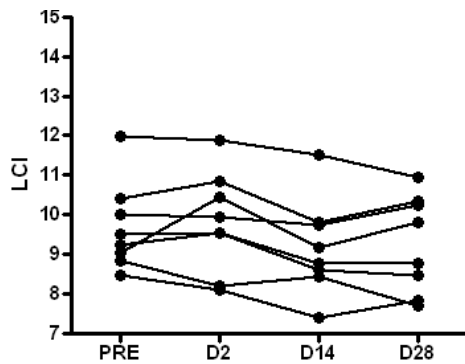


Figure 4.6A

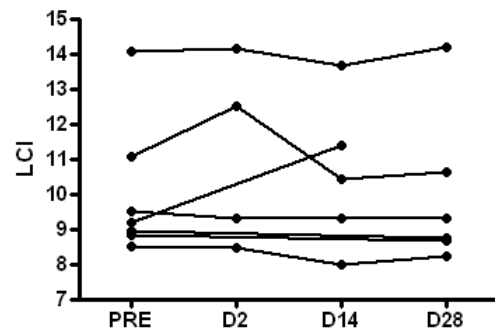


Figure 4.6B

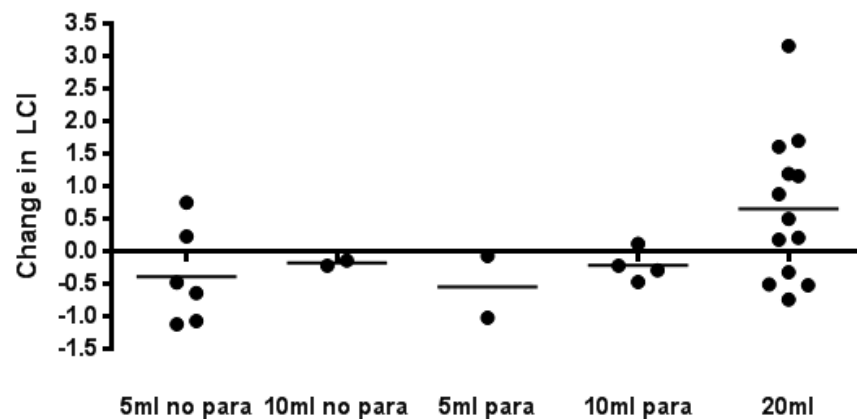


Figure 4.6C

Figure 4.6. Change in LCI according to dose cohort. Results for the 5ml lung dose cohort are in Figure 4.6A, and 10ml patients in 4.6B, with mean pre-dose LCI and LCI at each test date. Figures A&B are plotted irrespective of whether the patients received planned adjuvant paracetamol. All patients with a valid LCI test for both mean pre-dose and day 28 are summarised in Figure 4.6C. This is a plot of change from mean pre-dose LCI to day 28 result, the direction of improvement in LCI is towards negative values on the graph. This includes the separation of results according to whether the patients received planned adjuvant paracetamol (para = paracetamol). No 20ml patients received planned adjuvant treatment with paracetamol.

4.5.3 Inflammation

The post dose systemic inflammatory response as detected by laboratory markers (CRP, total white cell count (WCC)) is summarised in Figures 4.7 according to dose cohort. Results are presented for the maximal time points at which change was detected; for CRP this was day 2 post dose and for total WCC at 8 hours post dose.

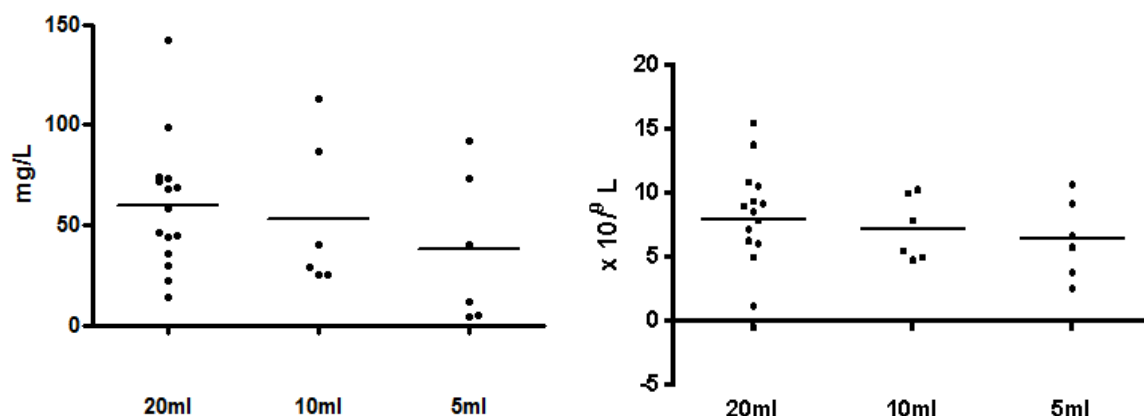


Figure 4.7A CRP result at day 2

4.7B Total WCC 8 hours post dose

Figure 4.7. Change in CRP and white cell count post dose. A) Maximum CRP post dose according to lung cohort on day 2 post dose. Although a trend was seen towards lower maximum values with lower volumes, the difference was not significant (Kruskal- Wallis $p=0.38$). B) Change in blood total white cell count from immediate pre-dose to 8 hours after the start of dosing. Similar trends were seen according to dose group as with CRP but the difference was non significant (KW $p=0.59$). Each patient is represented by an individual data point; those with planned adjuvant therapy are not included.

For all lung dose patients combined, there was no correlation between maximum CRP and maximum drop in FEV_1 (absolute %) post dosing ($r=0.22$, $p=0.19$) and similarly with change in total white cell count post dose ($r=0.3$, $p=0.07$). There was a significant correlation between the temperature measured 8 hours post dose and the maximal drop in FEV_1 (absolute %), $r=0.39$ and $p=0.02$.

4.5.4 Post dose follow up to day 28

Patients were seen for study visits according to protocol for their dosing cohort. In Pilot part A, ten patients that had had a pre-dose bronchoscopy underwent a subsequent post-dose bronchoscopy to look for molecular and functional evidence of successful gene transfer ($n=4$ on day 6 post dose; $n= 6$ at day 14), (see Chapter 6). Follow up for all patients in the Pilot study was until day 28, however additional visits were permitted. Reasons included further dates for nasal PD measurement ($n=4$) (for patients with on-going changes suggestive of efficacy at day 28, as described in Chapter 5), and pulmonary exacerbation coinciding with final scheduled trial visit ($n=1$).

4.6 Post hoc comparison of lung dose cohorts

To allow investigation of differences in safety outcome between dose groups, a post hoc analysis was undertaken to compare the characteristics of patients in each dosing cohort. There was no difference in enrolment FEV₁ % predicted between dose groups, or between age and dose group (Figure 4.8).

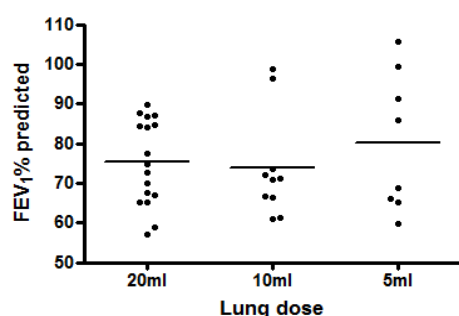


Figure 4.8A

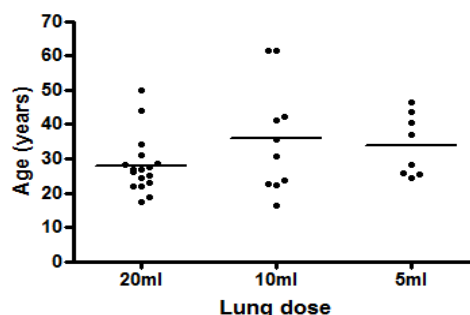


Figure 4.8B

Figure 4.8. FEV₁% predicted values and age at recruitment according to lung dosing cohort. Each patient is represented by an individual data point in Figure 4.8A for FEV₁% predicted and Figure 4.8B for age. There was no difference between groups for FEV₁% predicted or age (Kruskal Wallis test $p = 0.81$ and $p = 0.32$ respectively).

The baseline status in terms of usual management and co-morbidities of patients in each dose cohort is summarised in Table 4.2.

	Lung dose cohort		
	20ml (N=17)	10ml (N=10)	5ml (N=8)
Pancreatic insufficient (%)	88.2	80	75
Nebulised antibiotics (%)	82.4	70	87.5
RhDNase (%)	29.4	60	37.5
Hypertonic saline (%)	23.5	20	25
Long term azithromycin (%)	70.6	70	50
CF related diabetes on insulin (%)	23.5	10	12.5
Regular inhaled steroids (%)	35.3	70	37.5

Table 4.2. Characteristics of patients according to lung dosing cohort. Values in the table represent the % patients in each dose cohort with baseline status according to disease status or usual clinical management. Patients receiving planned adjuvant therapy have been included in the appropriate lung dose cohort and are not summarised separately.

4.7 Adjuvant therapy: 'As required' and planned

The majority of patients in pilot part A or pilot B without planned adjuvant therapy took paracetamol or ibuprofen on an 'as required' basis for symptomatic treatment post dosing (Table 4.3) although anti-pyretics were less commonly requested by patients in the 10 and 5 ml cohorts than by those receiving 20mls.

Lung dose	Paracetamol	Ibuprofen	Inhaled bronchodilators
20ml	13/15 (87%)	13/15 (87%)	10/15 (67%)
10ml	5/6 (83%)	1/6 (17%)	2/6 (33%)
5ml	3/6 (50%)	1/6 (17%)	2/6 (33%)

Table 4.3. Proportion of patients receiving additional medication on an 'as required' basis post dosing, according to lung dose cohort. This table excludes all patients with planned adjuvant treatment post dosing.

In total n=8 patients received *planned* adjuvant treatment with prednisolone (n=1), ibuprofen (n=1) or paracetamol (n=6). Patients receiving planned prednisolone or ibuprofen received a 20ml lung dose. The six patients receiving planned adjuvant therapy with regular paracetamol administration included n=4 patients receiving a 10ml dose and n=2 patients with a 5ml dose (none received the 20ml lung dose).

4.7.1 Ibuprofen

The post dosing profile for the patient receiving a 20ml lung dose with planned adjuvant ibuprofen was no different to the rest of the 20ml cohort. In addition to ibuprofen, he also received paracetamol for symptomatic relief in the immediate post dosing period. There was no evidence of a reduction in systemic inflammatory markers and a potential transient adverse impact on renal functioning meant that no further dosing with adjuvant ibuprofen was undertaken.

4.7.2 Prednisolone

In an alternative strategy to reduce the systemic inflammatory response, one patient received a 20ml lung dose with planned administration of prednisolone

as adjuvant therapy. This patient also received paracetamol and inhaled bronchodilators on an 'as required' basis in the period 48 hours post dosing. Again, lack of evidence for an impact on inflammatory response in combination with concern about the addition of prednisolone to the dosing schedule for any future multidose trial (including those relating to gene transfer efficacy and implications for placebo group) meant that this avenue was not pursued further.

4.7.3 Paracetamol

The post dosing profiles belonging to the n=6 patients receiving planned adjuvant treatment tended to have a lower rise in inflammatory markers post dose, although numbers are small for any statistical comparisons. No patients receiving this paracetamol dosing schedule requested 'as required' ibuprofen in the period 48 hours post dosing. The more favourable post dosing profiles observed in patients taking planned adjuvant paracetamol have led to its incorporation in the multidose trial for both active and placebo groups in an effort both to minimise adverse effects and aid with blinding.

4.7.4 Bronchodilators

All patients received pre-treatment with bronchodilators pre- nebulised gene therapy as per protocol. Overall, 16/35 (45.7%) patients received inhaled bronchodilators (either salbutamol or terbutaline) post dosing on an 'as-required' basis in the period up to 48 hours post dose (n=15 on dosing day, n=1 on day 2). Patients receiving a 20ml lung dose accounted for the majority of this group (n=11 for 20ml lung dose, n= 3 for 10ml lung dose, n= 2 for 5ml lung dose).

In patients taking planned adjuvant paracetamol or ibuprofen (total n=7), only one received additional 'as required' bronchodilators in the period 48 hours post dose (administered on day 2). However, from the data available it is not possible to determine whether this is a treatment effect of these agents or a reflection of the lower lung doses (10ml and 5ml only) administered to these patients. Patients reported little symptomatic improvement in response to bronchodilators and there was no obvious impact on spirometry.

4.8 Efficacy outcomes

4.8.1 CFTR molecular results

The mRNA results for the detection of vector specific *CFTR* from nasal brushings taken pre- and post dosing are summarised in Table 4.4. As anticipated, no pre-dose samples were positive for the vector specific CFTR mRNA. Overall, 21/22 patients had a post dose nasal brushing (one patient refused the test), the majority of which were also negative. Three post dose tests (14%) were positive for the presence of mRNA but in all cases the result was below the level deemed ‘quantifiable’. All positive results were obtained either at day 14 or 28, with none at earlier time points. Patient 030 had two brushings as was being followed up beyond the initial day 28 protocol due to evidence of ongoing chloride secretion on post dosing nasal PDs (Chapter 5). The mRNA results were negative on both occasions.

Patient	Nasal dose of gene therapy	Pre dose test taken?	Post dose test result	Day of post dose test
1SD001	1ml	N	Not detected	2
1SD002	1ml	N	Not detected	2
1SD004	1ml	N	Not detected	2
1SD003	2ml	Y	Not detected	2
1SD005	2ml	Y	Not detected	2
1SD020	2ml	N	Not detected	2
1SD014	2ml	Y	Not detected	6
1SD015	2ml	Y	Not detected	6
1SD023	2ml	Y	Not detected	6
1SD017	2ml	Y	Not detected	6
1SD010	2ml	Y	Not done	N/A
1SD021	2ml	Y	Positive: NQ	14
1SD022	2ml	Y	Not detected	14
1SD007	2ml	Y	Not detected	14
1SD011	2ml	Y	Positive: NQ	14
1SD012	2ml	Y	Not detected	14
1SD028	2ml	N	Positive: NQ	28
1SD029	2ml	N	Not detected	28
1SD030	2ml	N	Not detected	28 and 114
1SD031	2ml	N	Not detected	28
1SD034	2ml	N	Not detected	28
1SD041	2ml	N	Not detected	28

Table 4.4. Vector specific CFTR mRNA results from nasal brushings for all patients in the nasal cohort. For each patient, the nasal dose is shown along with whether there was a pre-dose sample taken (Y=yes, N=no). For post dose mRNA results,

Not detected= no vector specific CFTR mRNA detected; Positive: NQ= detected but not quantifiable.

Similarly the mRNA results from lower airway brushings taken during the pre- and post- dose bronchoscopy are summarised in Table 4.5. Post dose bronchial brushings were obtained on day 6 (n=4) and day 14 (n=6). Again the majority of post dose measurements were negative, with only 2 patients (20%) having vector specific *CFTR* mRNA detected.

Patient	Lung dose of gene therapy	Pre dose test?	Post dose test result	Day of post dose test
1SD014	20ml	No data	Not detected	6
1SD015	20ml	Y	Not detected	6
1SD023	20ml	Y	Positive: NQ	6
1SD017	20ml	Y	Not detected	6
1SD010	20ml	Y	Not detected	14
1SD021	20ml	Y	Not detected	14
1SD022	20ml	Y	Not detected	14
1SD007	20ml	Y	Positive: 2.39%	14
1SD011	20ml	Y	Not detected	14
1SD012	20ml	Y	Not detected	14

Table 4.5. mRNA results from bronchial brushings for all patients in the bronchoscopic cohort (Pilot part A). For each patient, the lung dose is shown along with whether a pre-dose brushing was tested (Y=yes, No data=no sample analysed). For post dose mRNA results, ND= no vector specific CFTR mRNA detected; Positive: NQ= detected but not quantifiable. Not detected: no vector specific CFTR mRNA detected; NQ: not quantifiable. One patient had a positive quantifiable result, with the value (2.39%) representing the proportion of vector specific CFTR mRNA copies / human CFTR mRNA copies by TaqMan RT-PCR assay

Patients in the bronchoscopic cohort had their post dose nasal brushing on the same date. No patient had a positive mRNA result on both nasal and bronchial brushings. Comparisons between molecular evidence of efficacy and the results of functional assessment of CFTR (airway PD measurements) are described in Chapter 5 (nasal PD) and Chapter 6 (bronchial PD) respectively.

4.9 Serious adverse events

Two serious adverse events (SAEs) was reported during the trial (one was reported as a suspected unexpected serious adverse reaction (SUSAR), see

below). One SAE was reported following a pre dose bronchoscopy. This patient experienced severe throat pain post bronchoscopy which, on examination was found to be the result of a pressure area on the uvula from a malpositioned endotracheal tube. He was observed overnight in hospital in case of airway compromise however remained well and was discharged the following morning. No intervention except simple analgesia was required. He was discharged home the following morning and agreed to undergo subsequent post-dosing bronchoscopy as per protocol, which was uneventful.

One patient in the first 10ml nebulised/1ml nasal cohort, who had been well at a day 6 follow up visit, was admitted to a local hospital for acute pancreatitis prior to his day 14 follow up visit (which he therefore did not attend). Although we considered it unlikely to be related to the administered gene therapy product, this event was reported as a SUSAR. Despite never having been diagnosed formally as having a history of pancreatitis, we subsequently noted that he had had several previous admissions for abdominal pain and vomiting over the preceding years; on none of these occasions was a serum amylase measured. Following this, serum amylase levels from all other 10ml patients (and all subsequent 20ml trial patients) were measured and all were unremarkable.

4.10 Discussion

The Pilot study was undertaken to assess the safety and functional efficacy of the CF gene therapy product pGM169/GL67A. The evolution of the Trial protocol in response to safety outcomes reflected the ‘first-in-man’ nature of the trial and the complexities associated with evaluating outcome and determining tolerated delivered dose. This process has also allowed an investigation of the outcome measures in the Pilot study, and has raised a number of themes for discussion.

4.10.1 Trial design

4.10.1.1 Recruitment

The Pilot study included a large number of outcome measures, in part reflecting the dual purpose of assessing safety and functional efficacy. Prior knowledge of the most appropriate time points for data collection (for either safety or efficacy measures) would have allowed tailoring of the trial protocol at the design stage to reflect this. However, this is not necessarily possible with early phase clinical trials. Guidance was sought from pre-clinical studies and the previous Lancet study (35). However, the gene therapy plasmid was different (having been optimised for duration of expression and removal of CpG repeats as described in Chapter 2, section 2.2.1) and so the time points appropriate for efficacy in the Lancet study were not necessarily appropriate for Pilot. Furthermore, mice are not human: the administration of pGM169/GL67A to mice in a pre-clinical study prior to trial onset did not result in a systemic inflammatory response (100).

The Pilot post dosing schedule, particularly in part A, required a high level of commitment from participants both in terms of number and length of study visits. This may explain the relatively large number of patients enrolled in the study withdrawing prior to dosing (section 4.1).

4.10.1.2 Inclusion criteria and pre-dosing visits

As the Pilot study was primarily a safety study and not designed to test clinical efficacy, the concept of ceiling effects for outcome measures described in the Run-In discussion (Chapter 3, section 3.7.6) is less relevant. There was no maximum cut off for lung function in terms of % predicted FEV₁ at recruitment, however studies investigating clinical efficacy using FEV₁ as a primary outcome measure would need to consider this. It is important that safety studies do not recruit a patient population considered 'unsuitable' for subsequent efficacy trials to ensure results are representative. Therefore it was appropriate that the minimal FEV₁% predicted at enrolment for the Pilot study was 10% lower than for the Lancet study (35).

4.10.1.3 Number of pre-dose time points

The Pilot study design was strengthened by collection of serial measurements for certain outcome measures (such as spirometry, LCI, blood markers and nasal PD) pre- gene therapy dose. The Run-In data (Chapter 3) has described the variability of airway physiology tests and systemic markers of inflammation over time (Chapter 3, section 3.3). However, for other assays, only a single time point of data collection existed pre dose. This reflected either trial design (e.g. single measurement only for gas transfer measurement pre dose), or significant additional trial burden with incorporation of a second pre-dose bronchoscopy or CT scan (including potential excess morbidity from general anaesthetic and radiation exposure). Whether comparisons should be made according to a single (e.g. initial visit) or mean pre-dose values is important as it can influence the magnitude of any effect size. For example, the results profile for LCI data in Pilot could appear significantly 'improved' by comparing only initial pre-dose LCI with post dose measurements, rather than mean pre- values. Therefore in the results described within this chapter only mean pre dose results were shown for comparison with post dose results (Figure 4.6).

Multiple pre-intervention study visits have not commonly been incorporated into clinical trial protocols for CF gene therapy or other novel treatments. However, this may be important to consider depending on the study populations, particularly if variability in outcome measures is increased with greater severity of underlying disease as suggested by analysis of the Run-In airway physiology and patient reported outcomes (Chapter 3, section 3.3.5).

4.10.1.4 Dose administration

Reference to dose cohorts according to volume of the lung dose of gene therapy (20ml, 10ml etc.) has related to the intended dose prescribed. However, weighing of the nebuliser and nasal spray bottles demonstrated that this was not equivalent to the actual dose administered (section 4.3). Reassuringly, the proportion of actual administered dose was equivalent for the 20ml, 10ml and 5ml cohorts. The variability of dose received was lower for lung than nasal dosing (as described previously) and within-dose cohort differences in volumes were minimal. However, any differences in breathing technique may have led to hypothetical discrepancies in the amount of product reaching the distal airways.

The majority of patients were used to taking at least one nebulised therapy (antibiotics, RhDNase or hypertonic saline) as part of their usual CF care and therefore the concept of a nebulised therapy was not new.

It was surprising that the volume of gene therapy delivered using the nasal spray bottle was so variable between patients (Figure 4.2), although patient technique and method of pressing the nasal spray bottle mechanism, despite supervision, may be responsible for this. Of particular concern was the patient who was prescribed a 2ml nasal dose whose actual delivered dose was 0.97ml, representing significant underdosing in comparison to intended delivery. To minimise any influence of unintentional dose-response differences, it is important to ensure that an intended volume is delivered consistently between patients, particularly with repeated dosing (as occurs in the multidose trial) as such dose discrepancies may assume a greater importance. This could easily be achieved by weighing nasal spray bottles (and/or nebuliser chambers) during dosing to guide volumes administered rather than at conclusion of dosing as has been carried out in the Pilot study. This will then ensure that comparison between different outcome measures is dependent on 'true' response rather than variability in administered dose.

4.10.2 *Safety outcomes*

4.10.2.1 Relationship to lung dose cohort

We observed a post dose inflammatory response following administration of the lung dose (section 4.5). The magnitude of this response was dose dependent, with patients receiving the 20ml lung dose experiencing greater adverse response than those in the 5 or 10ml cohort. The outcomes collected for spirometry, LCI and gas transfer provide information on different aspects of airway function. They collectively capture a cross-sectional snapshot of function at a particular time point. The sites of importance in terms of disease pathology (e.g. small airways and information provided by LCI) may be useful sites to measure for both safety and efficacy. Conversely, others may only be important for safety (e.g. gas transfer) in the relatively well CF populations included in the UKCFGTC clinical programme.

4.10.2.2 Airway physiology

The post dose response following nebulised gene therapy was associated with a decline in FEV₁ and FVC in the 6 hours post dosing. Repeated spirometry measurements during this time were often met with reluctance by the patient. Knowledge of the time period over which this decline occurs now prevents the requirement for such intense monitoring during the multidose trial. A single measurement at e.g. 6 hours post dose is also likely to be more representative of ‘true’ lung function at that point rather than also reflecting fatigue.

Gas transfer is a test of alveolar function (rather than that of the conducting airways as assessed by spirometry and lung clearance index). It is not usually impaired in cystic fibrosis lung disease, as it is the small airways rather than alveoli which contribute to greatest morbidity. However, its inclusion in the Pilot protocol reflects its role as a safety outcome. Any new intervention such as a nebulised gene therapy product may impact on mechanisms outside the traditional pathways of disease pathogenesis associated with cystic fibrosis and therefore tests are required that can detect change reflecting this.

We observed a decrease in gas transfer at day 2 which had normalised by day 14, again with a greater decline seen in the 20ml dose cohort (section 4.5.2). Gas transfer was not measured in the period 6 hours post dose (the time point of maximal change in spirometry) nor on day 1 post dose. It is therefore unknown whether the results at day 2 represent an improvement from an earlier time point, or a genuine lag behind the more immediate decline in FEV₁ and FVC. Comparisons between pre- and post dose gas transfer measurements relied on a single pre-dose test. As demonstrated with other physiological outcomes in the Run-In study (Chapter 3, section 3.3), there is a variability over time which is important to consider when assessing change in response to an intervention.

4.10.2.3 Systemic inflammatory response

Counter to our expectations based on plasmid design and pre-clinical animal testing, it became clear early in the trial that a familiar flu-like response was being experienced during the 48 hours post dosing. Given the complete removal of all CpGs from the plasmid, the pathological processes driving this response remain uncertain; it is possible that some genomic unmethylated CpG repeats

exist in the preparation and work is ongoing within the UKCFGTC to identify the source of this.

Whilst the severity of the inflammatory response in terms of symptoms and biochemical/clinical parameters appeared to be related to dose on a group level (section 4.5), for individual patients responses were highly variable. In addition to less severe adverse responses in the 10ml and 5ml lung dose patients in comparison to those receiving 20mls, they also requested less ‘as required’ medications for adverse symptoms (section 4.7, Table 4.3). Therefore the greater tolerability of the lower doses was likely to have been genuine.

Detailed profiling of laboratory markers and clinical outcomes has identified the chronology of the response, in order that further monitoring in a multi-dose trial can be rationalised similar to that described for repeated spirometry results as described above.

4.10.3 Efficacy

4.10.3.1 Molecular evidence of effective gene transfer

The rate of positive detection for vector specific *CFTR* mRNA in nasal and lower airway brushings was low (section 4.8.1). An absence of association between upper and lower airway results, and with functional measurements in the upper and lower airway (Chapter 6) suggests that this relates to low sensitivity of the assay, but interpretation is difficult. However, despite these results, there is encouraging molecular data from researchers within the Consortium at the Oxford site following repeated dosing with pGM169/GL67A in mice (unpublished). Therefore, assessment of mRNA may be more appropriate as an endpoint in trials following repeated rather than single occasion dosing. However, if this is the case, the results from the Pilot study will provide important comparative data at an early time point, assuming the patient populations are representative of each other.

4.10.3.2 Clinical end points

The Pilot study was not designed (or powered) to assess clinical efficacy; indeed, we did not expect clinical benefit after a single dose of gene therapy.

Therefore, it was surprising to observe the changes in LCI in the 5 and 10 ml lung group (Figure 4.6) as LCI was incorporated as a safety outcome, and it is difficult to know how best to interpret these. Unfortunately, data quality at trial onset for LCI meant that a proportion of early tests had to be excluded (Chapter 2, section 2.3.2.11). Reasons for this have been discussed in Chapter 3, section 3.7.1. However, despite these limitations, the LCI data is nevertheless of interest and LCI has been included as a major secondary outcome in the subsequent multidose trial. LCI was not measured in the Lancet study (or any other previous trial of gene therapy) and therefore no comparisons can be made. Although as an outcome measure for CF LCI remains in its infancy, detectable changes in treatment groups post intervention have been reported. This includes improvement following hypertonic saline, RhDNase and the small molecule drug Ivacaftor (50;69;70). These studies all reported LCI as an outcome measure in terms of absolute change. In the Pilot study, it would be inappropriate to map any change in LCI suggestive of improvement with change in other outcomes such as FEV₁, however following the multidose trial this will be possible. An increasing body of literature incorporating LCI as an outcome measure in CF will help determine appropriate minimal clinical differences and whether this should be considered as absolute change from baseline, percentage change, or change in relation to normal values according to age and gender. It may be that a combined small improvement in several outcomes (e.g. both FEV₁ and LCI) is of equivalent clinical significance to a larger improvement in one alone (FEV₁). Again, Pilot study data is inappropriate for such an analysis and the results from multidose are awaited.

4.10.4 *Adjuvant therapy*

Surprisingly, planned adjuvant therapy with ibuprofen or prednisolone did not alter the typical post dose profile. However, paracetamol did appear to be beneficial, measured both by a dampening of the systemic markers of inflammation and by patient report of influenza-like symptoms (less additional ‘as required’ additional adjuvant therapy was taken by this group). There is evidence that paracetamol is effective at reducing influenza-induced pathology (including decreased infiltration of inflammatory cells into the airway spaces and preserving lung function) in mice (122).

The use of any adjuvant therapy in a clinical trial creates additional challenges for data analysis and appropriate placebo in subsequent multidose trial. However its absence would create problems with blinding both for the patient and researcher. Inclusion of planned adjuvant treatment with paracetamol in both active treatment and placebo groups in the multidose trial is hoped to address this.

4.10.5 *Conclusions*

This chapter has discussed the results of safety outcomes including airway physiology, clinical and laboratory markers of inflammation and molecular tests of functional efficacy. In conjunction with the results from airway PD outcomes which are described in Chapters 5 and 6, this has allowed an assessment of the suitability of pGM169/GL67A to be taken forward to multidose trial according to review by the CF Trust's Scientific Advisory Committee, and to rationalise protocol design to reflect appropriate time points of testing. It is always going to be a challenge to design a clinical trial for a novel therapeutic agent which has not previously been administered to man. Difficulties exist with small 'n' numbers, choosing appropriate markers for comparisons (with variables introduced by factors to do with delivered dose and time for administration) and including single or multiple pre-intervention measurements balanced against an acceptable trial protocol from the perspective of the patients volunteering to take part.

Despite these challenges, the Pilot study was able to quantify differences in the safety profile according to delivered dose and thus able to select the 5ml dose to take forward to the multi-dose trial.

5 Results: nasal PD measurement

5.1 Pilot study nasal PDs

The total number of nasal PD tests completed in the Pilot study was 143 in 33 subjects. An additional 4 tests (involving 3 subjects) were abandoned part way through due to patient discomfort (n=2 tests (same subject, 021)), or inability to maintain a stable baseline (n=2 tests)(see Table 5.1 and below). All within-subject measurements throughout the study were performed in the same nostril with the exception of three tests where the other nostril was used (inability to achieve a basal PD measurement (n=1; subject 002), uncharacteristic response to Ringers with amiloride perfusion (n=1; subject 011) and pain on insertion of PD catheter (n=1; subject 021 as above)). In all 3 subjects, measurement in the other nostril occurred on a single occasion only therefore in none did a problem exist where all pre-dose tests were performed on a different side to post-dose tests.

Eleven patients had at least one nasal PD attempted but did not go on to receive a nasal dose (total n=17 tests, including one of the abandoned tests). Reasons for this included evidence of chloride secretion on nasal PD in pre-treatment measurements (n=5 subjects), and withdrawal from study pre-dosing either due to patient request or clinical instability (n=6).

Twenty two subjects received a nasal dose of gene therapy, with three receiving a 1ml nasal dose and 19 a 2ml nasal dose as described in Chapter 4. The schedule of nasal PD tests for the 127 tests undertaken within patients receiving a nasal dose is summarised in Table 5.1.

Patient	Nasal dose	No. Pre-dose tests	Day 2	Day 6	Day 14	Day 28	> Day 28
001	1ml	2	Y				
002	1ml	2	Y				
003	2ml	3	Y				
004	1ml	2	Y				
005	2ml	3	Y				
007	2ml	2	Y		Y	Y	Y (1)
010	2ml	3	Y	Y	Y	Y	
011	2ml	3	Y	Y	Y	Y	
012	2ml	3	Y	Y	Y	Y	
014	2ml	3	Y	Y	Y		
015	2ml	3	Y	Y	Y		
017	2ml	3	Y	Y		Y	
020	2ml	2	Y				
021	2ml	2*	Y		*		
022	2ml	3	Y	Y	Y	Y	Y (3)
023	2ml	3	Y	Y	Y		
028	2ml	3	Y	Y	Y	Y	
029	2ml	3		Y	Y	Y	
030	2ml	3		Y	Y	Y	Y (5)*
031	2ml	3		Y	Y	Y	
034	2ml	3		Y	Y	Y	
041	2ml	3		Y	Y	Y	Y (1)
Total		60	17	14	14	12	10

Table 5.1. Schedule of nasal PDs for each patient receiving a 1ml or 2ml dose of CF nasal gene therapy. For tests after day 28 (> Day 28), the number represents the number of additional tests performed.*Denotes an additional test which was intended but not completed. For 021, the reason included inability to tolerate the PD catheter, and for patient 030 an intercurrent upper respiratory tract infection meant that testing was not possible.

The number of nasal PDs undertaken in an individual patient receiving the nasal dose of CF gene therapy ranged from 3 to 11. Four subjects returned for at least one visit after day 28 for a nasal PD, at intervals convenient to the patient and depending on the visual analysis of the PD trace. One of these patients was subject 030, who also had a nasal PD performed opportunistically as part of another study during this period of extended follow up. This test is not included in the above table, nor are formal results reported here (due to differences in the experimental set up) however evidence of chloride secretion on this test led to a further Pilot nasal PD for which the results are described below.

5.2 Pilot study nasal PD results

5.2.1 Nasal PD circuit integrity

The mean offset for electrodes prior to offset correction across all nasal PD tests was -0.55mV (SD 2.56 mV). Skin PD was greater (more negative) than -30mV for each patient on each test occasion.

5.2.2 Agreement between analysers

The outcome of the consensus review process between GD and JD for all 118 nasal PD tests from patients with a 2ml nasal dose (472 sections for inter-reporter agreement analysis) is summarised in the flow chart in Figure 5.1.

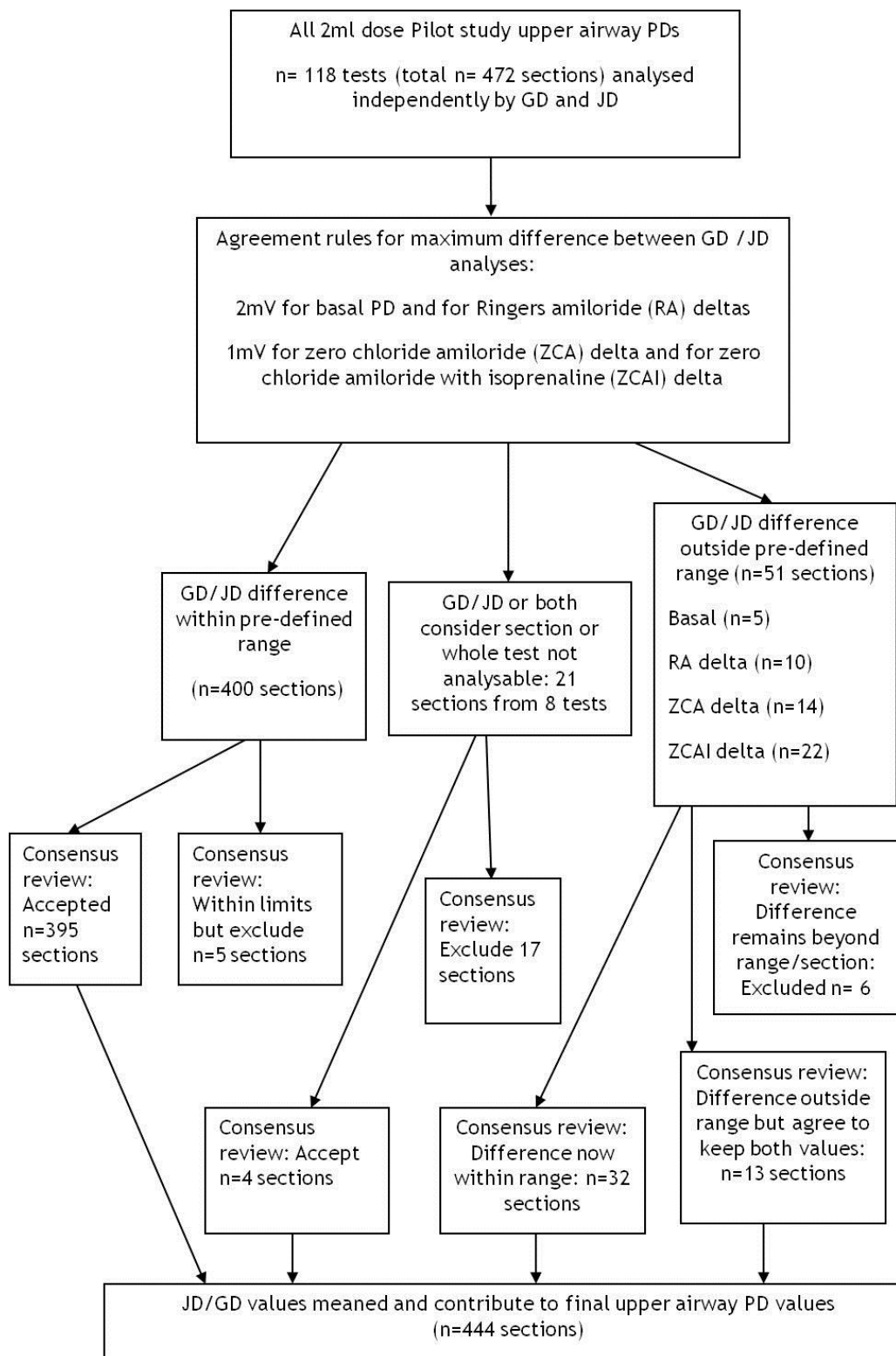


Figure 5.1. Flow chart describing outcome of nasal PD data analysis and consensus review process between GD and JD. RA: ringers with amiloride, ZCA: zero chloride with amiloride, ZCAI: zero chloride with amiloride and isoprenaline. For the n=13 sections surviving the consensus review process but remaining outside the initial ‘rules’ for maximum differences, the magnitude of the difference was as follows: basal PD (n=1) 2.001mV; ZCA (n=3): mean 1.7, max 2.05mV; and ZCAI (n=9): mean 1.5, max 2.50mV).

There was no directional bias to the analysis for nasal PD values pre-consensus between JD and GD (the Bland Altman plots can be found in Appendix 9.6). Following comparison of the independently analysed results, JD and GD reviewed tests where a discrepancy existed between analysers, or which either or both analysers felt required further discussion. This resulted in several sections of tests being excluded from data analysis. Reasons for tests not being analysable included electrical artefacts during testing or multiple steep ‘blips’ in the mV trace unlikely to reflect underlying ion transport. Several tests had an interpretable baseline but the zero chloride response was not interpretable.

395/472 (84%) sections of nasal PD tests passed the GD/JD initial inter-reporter agreement process of being within pre-defined maximal differences for each test section and did not require further review. For all other sections, following discussion about inclusion or exclusion (including those with a mV difference outwith the maximum pre-defined difference for each variable), repeat analysis was performed by one or both persons, still blinded as to identifying details. This resulted in 49 additional acceptable sections (10%) to define the final PD values. These included 13 sections where the difference remained outside the maximal difference following the review process, but the consensus from both GD and JD was to include both values to contribute to the final mean value (as each was considered equally valid). The remaining 28 sections (6%) were excluded from contributing to final PD results following review by GD and JD. Therefore following the review process, 444 sections (representing 94.1% of total sections) of nasal PDs were included in the final dataset. The Intraclass correlation coefficients (ICCs) for results for these sections prior to and following the consensing exercise are summarised in Table 5.2.

	Basal PD	RA delta	ZCA delta	ZCAI delta
Independent analysis (initial)	0.997	0.993	0.954	0.899
Final values (following consensus review)	0.999	0.998	0.997	0.986

Table 5.2. Intra-reporter variability in nasal PD analysis. The values in the table are the intra class correlation coefficients (ICC) for GD/JD nasal PD analysis for all patients with 2ml nasal dose for initial independent analysis (top row) and following consensus review (bottom row). RA: ringers with amiloride, ZCA: zero chloride with amiloride, ZCAI: zero chloride with amiloride and isoprenaline.

5.2.3 Nasal PD results for initial (1ml) nasal dose cohort

The pre-dosing results from the three patients receiving a 1ml nasal dose are summarised in Table 5.3. The SD for pre-dose basal PD in subject 002 was large - each of the two pre-dose tests were performed in a different nostril as a stable baseline could not be achieved in the other nostril, as described above.

Patient	Mean basal	SD basal	Mean RA delta	SD RA	Mean ZCA delta	SD ZCA	Mean ZCAI delta	SD ZCAI
001	-45.0	0.2	-21.0	4.4	5.9	3.4	-2.5	2.1
002	-36.6	23.1	-17.6	18.0	-3.2	3.4	-3.8	4.1
004	-44.9	2.8	-32.2	0.4	-2.8	4.6	4.8*	N/A

Table 5.3. Pre-dosing nasal PD summary for the 1ml nasal dose patients. The values in the table (in millivolts) are mean and standard deviations (SD) for each section of nasal PD, for all pre-dosing tests. *Represents single measurement only. RA: ringers with amiloride, ZCA: zero chloride with amiloride, ZCAI: zero chloride with amiloride and isoprenaline.

Post treatment nasal PDs in the first 3 patients involved in the Pilot study (1ml nasal dose) were limited to day 2 post treatment. This coincided with the timing of the post dose nasal brushing and no further nasal PDs were attempted after this time point, as per protocol design. Results for intrinsic zero chloride response (ZCA delta) and total chloride secretion (ZCA + ZCAI delta) are shown in Figures 5.2 and 5.3 respectively.

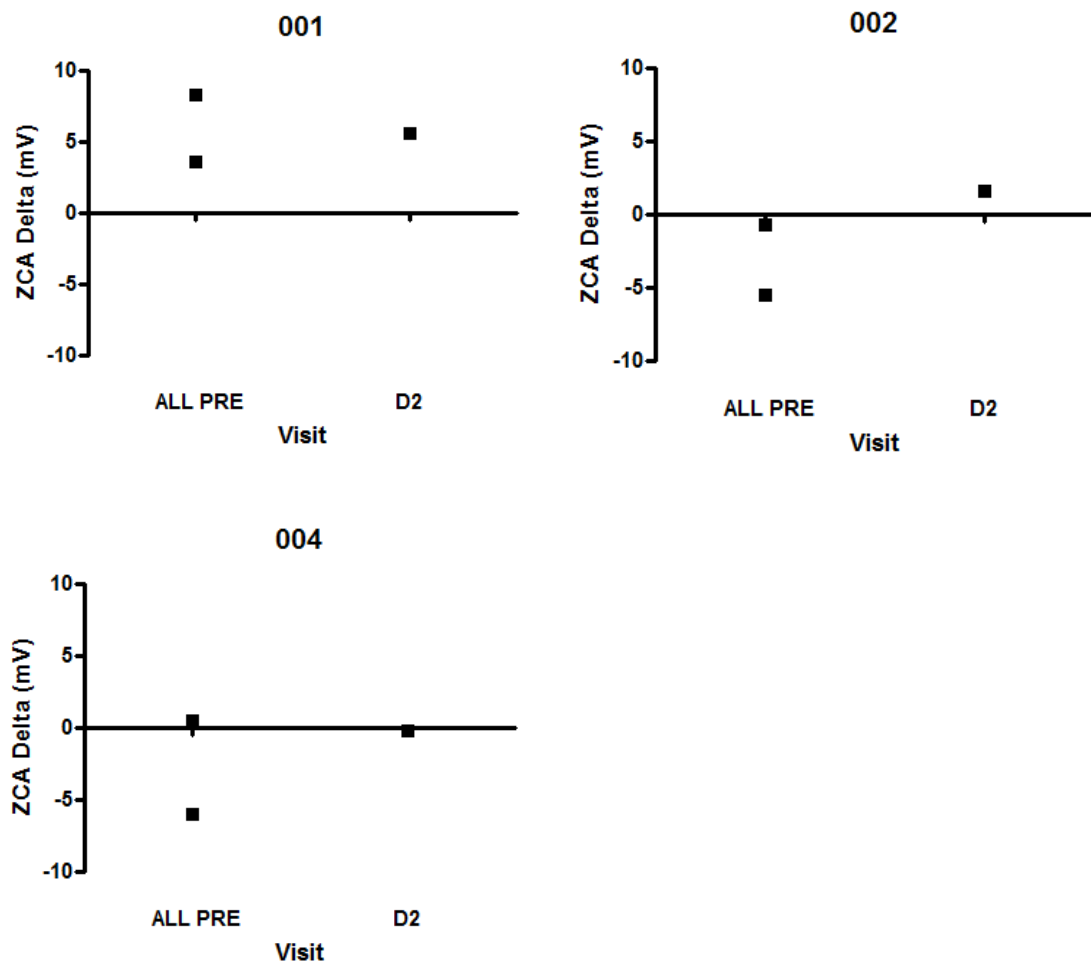


Figure 5.2. Intrinsic chloride secretion at all pre- and post dosing tests for the n=3 patients receiving a 1ml nasal dose. Patients identified according to study number (001, 002 and 004). ZCA: Zero chloride with amiloride; D2: day 2 post nasal gene therapy dose.

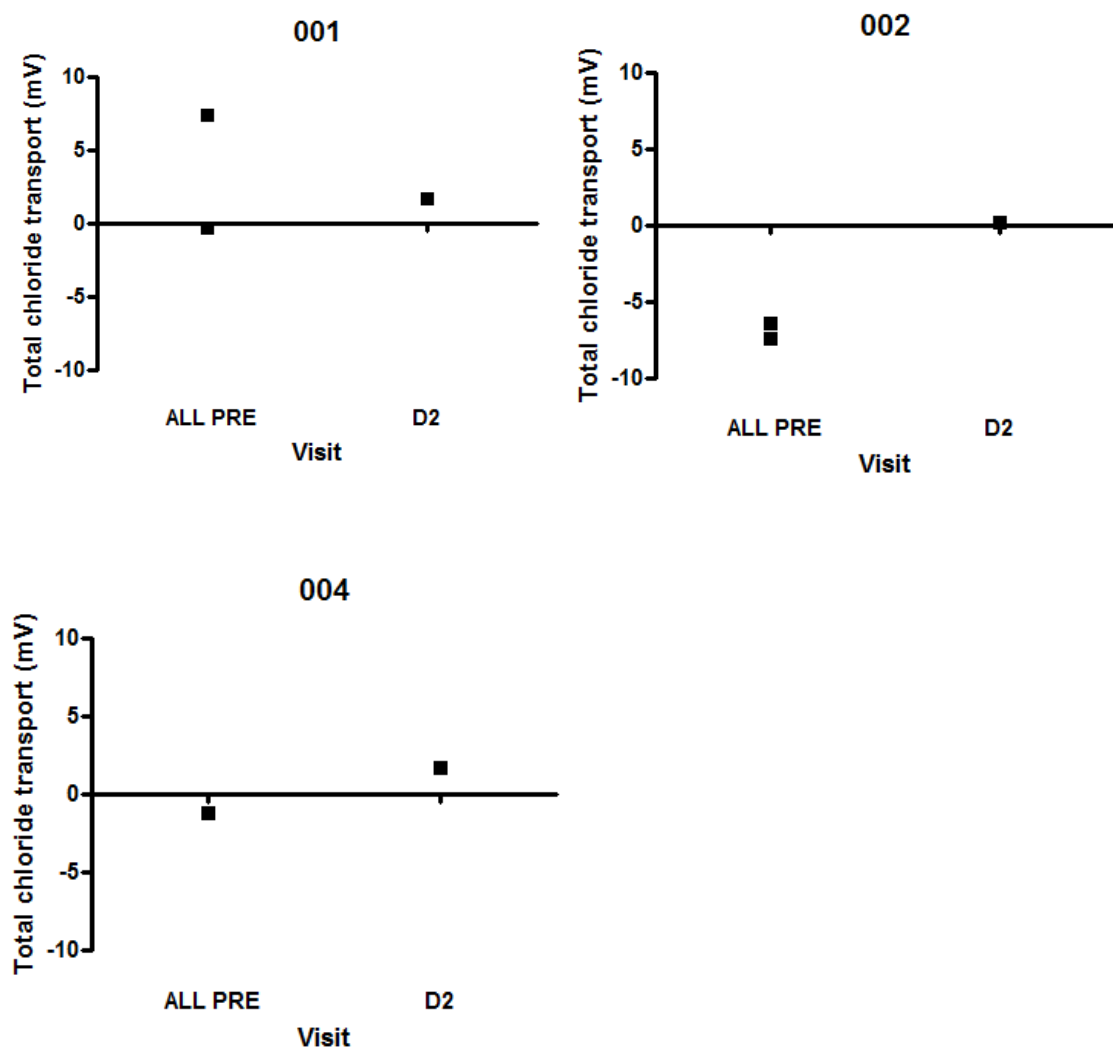


Figure 5.3. Total chloride secretion at all pre- and post dosing tests for the n=3 patients receiving a 1ml nasal dose. Patients identified according to study number (001, 002 and 004) as before. Total chloride secretion is the sum of deltas for zero chloride with amiloride and zero chloride with amiloride and isoprenaline response. D2: day 2 post nasal gene therapy dose.

5.2.4 Final PD values for main (2ml) nasal dose cohort

All results for the n=19 patients receiving the 2ml nasal dose are the final GD/JD values following the consensus review process for each section of the PD test (basal, delta RA, ZCA and ZCAI responses).

5.2.4.1 Variability in pre-dosing PD measurements

All 19 subjects had at least two pre-dose nasal PD tests available for analysis, and 16/19 (85%) had 3 pre-dosing PDs. However one patient (patient 003) with three pre-dosing tests only had two with values contributing to final results as their entire initial visit PD was excluded as it had been deemed unanalysable.

The pre-gene therapy nasal PD results for basal measurements and deltas for RA, ZCA and ZCAI sections are displayed in Table 5.4, with mean and standard deviation (SD) displayed for each patient. Overall for the entire 2ml nasal dose cohort, the median coefficient of variation for basal PD was 17.1%, and 32.2% for the RA delta. Due to the proximity to zero and the presence of negative and positive values for chloride transport the CV is invalid so variability is summarised by SD alone.

Patient	Mean basal	SD basal	Mean RA delta	SD RA	Mean ZCA delta	SD ZCA	Mean ZCAI delta	SD ZCAI
028	-37.2	9.6	-8.5	2.2	1.8	0.7	-1.3	1.1
029	-52.1	10.3	-27.3	24.3	-1.3	0.1	-0.5	4.0
030	-21.4	2.5	-8.2	2.6	0.3	0.8	-0.4	2.0
031	-35.8	6.2	-23.2	9.9	-1.9	2.2	-0.9	0.4
034	-31.9	2.5	-19.8	5.5	-1.9	1.2	0.2	0.6
041	-46.4	15.8	-25.9	3.3	-0.7	3.1	0.7	N/A
003	-40.3	5.8	-17.9	1.4	-5.0	1.4	-0.7	1.1
005	-65.0	4.5	-42.9	8.4	-6.6	7.6	-3.0	4.4
007	-36.2	10.0	-18.9	4.1	-1.5	3.4	-2.0	N/A
010	-38.5	6.6	-16.8	8.0	3.2	0.5	-0.3	2.7
011	-42.0	6.0	-12.5	7.0	-2.0	3.3	-2.3	1.8
012	-50.1	16.7	-22.8	11.3	0.8	1.4	-4.5	3.0
014	-49.8	4.7	-20.1	7.5	-0.8	0.2	-3.7	1.2
015	-21.1	1.8	-6.3	3.6	0.4	1.6	-1.1	1.1
017	-51.4	11.1	-24.1	8.1	-0.7	1.0	1.5	2.4
020	-48.3	5.3	-21.9	2.6	5.3	0.5	-1.8	1.7
021	-37.1	4.0	-13.3	1.7	1.4	N/A	-2.7	N/A
022	-29.9	6.5	-12.6	9.6	-2.8	2.8	0.0	1.8
023	-41.7	20.5	-17.5	3.3	0.8	3.5	-2.7	1.2
Mean	-40.9	7.9	-19.0	6.6	-0.6	2.0	-1.3	1.9
Median	-40.3	6.2	-18.9	5.5	-0.7	1.4	-1.1	1.8

Table 5.4. Pre-dosing nasal PD summary for the 2ml nasal dose patients. The values in the table (in millivolts) are mean and standard deviation (SD) results for each section of nasal PD, for all pre-dosing tests. Missing values are present when only one pre-dose result was available. RA: ringers with amiloride, ZCA: zero chloride with amiloride, ZCAI: zero chloride with amiloride and isoprenaline.

5.2.4.2 Evidence of CFTR correction and timing of nasal PD

Comparisons between pre- and post-dosing measurements were based on the mean pre-dose values as described above in Table 5.4.

The number of nasal PDs performed at post dosing study visits was n=14 at day 2, 6, and 14 and n=12 at day 28. The stage of the Pilot study (A or B) at dosing for each patient determined the dates of PD tests. For example, for some patients

the latest post dose test was at day 2 (e.g. initial non-bronchoscopic arm in Pilot A), whereas for others later in the trial nasal PDs were performed until day 28 visit.

5.2.4.3 Basal PD and response to perfusion with Ringers with amiloride solution

Basal and Ringers amiloride responses were analysed to investigate the influence of the single dose of gene therapy on sodium hyperabsorption. No differences were seen between basal PD pre- and post-dosing measurements. The paired results for mean pre- basal PD and post dosing measurements at day 2,6,14 and 28 are shown in Figure 5.4A-D. Similarly, no differences were seen in a direction consistent with non-CF values between mean pre- and post-dosing measurements in the change in PD as a result of perfusion with Ringers containing amiloride (Figure 5.4E-H).

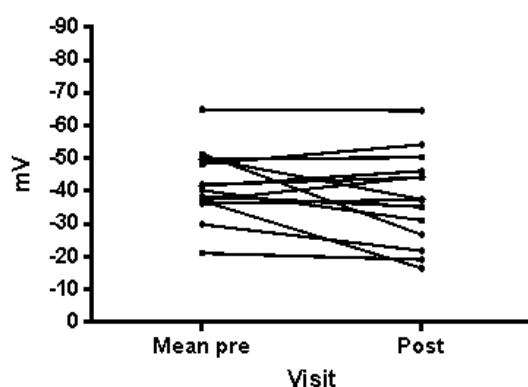


Figure 5.4A. Basal to day 2

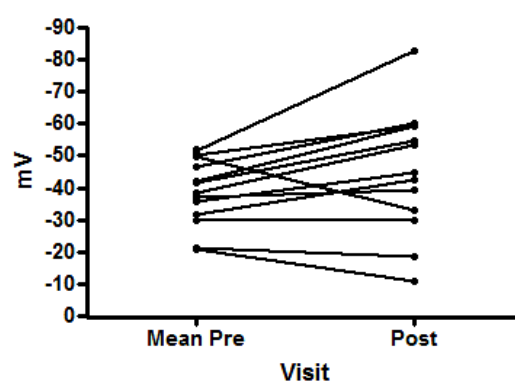


Figure 5.4B. Basal to day 6

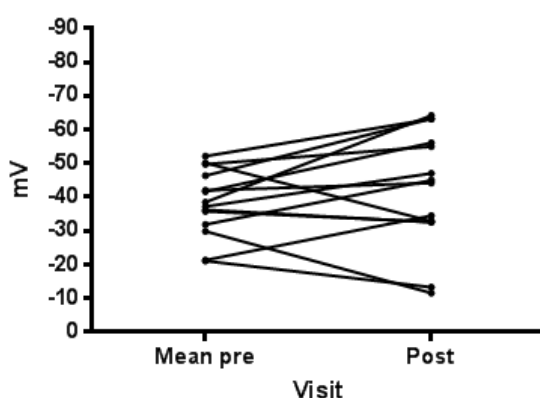


Figure 5.4C. Basal to day 14

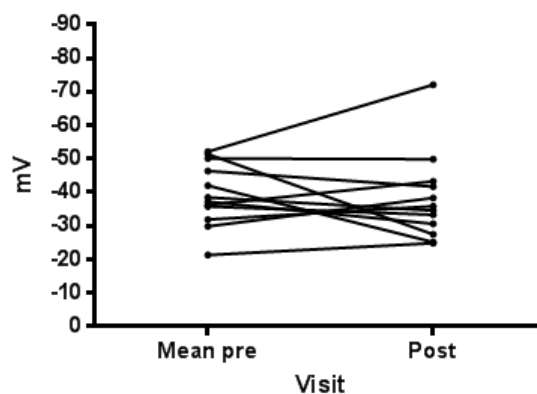


Figure 5.4D. Basal to day 28

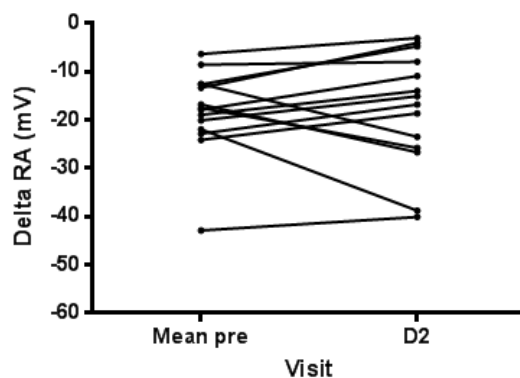


Figure 5.4E. Delta RA to day 2

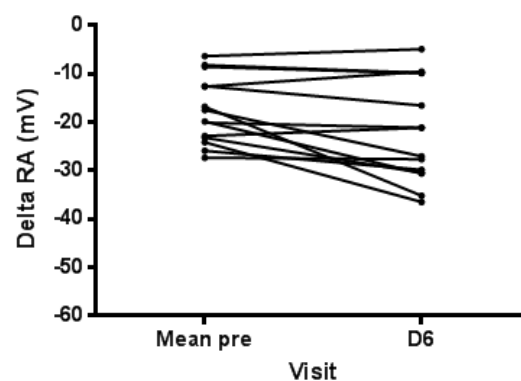


Figure 5.4F. Delta RA to day 6

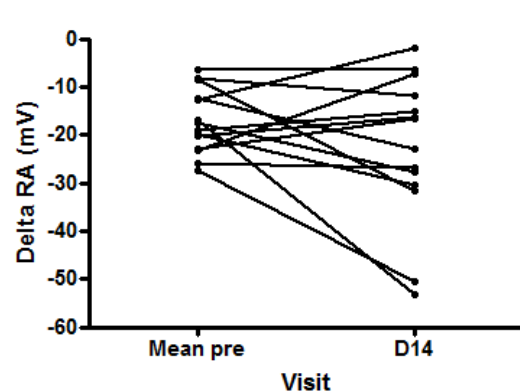


Figure 5.4G. Delta RA to day 14

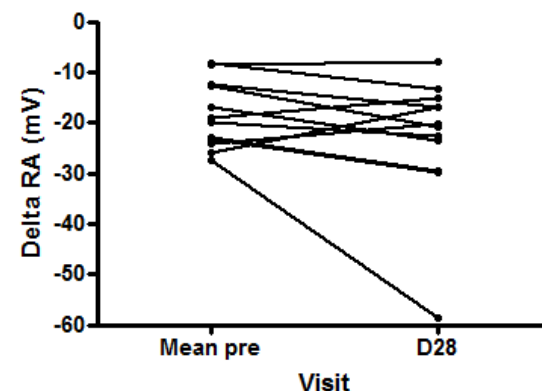


Figure 5.4H. Delta RA to day 28

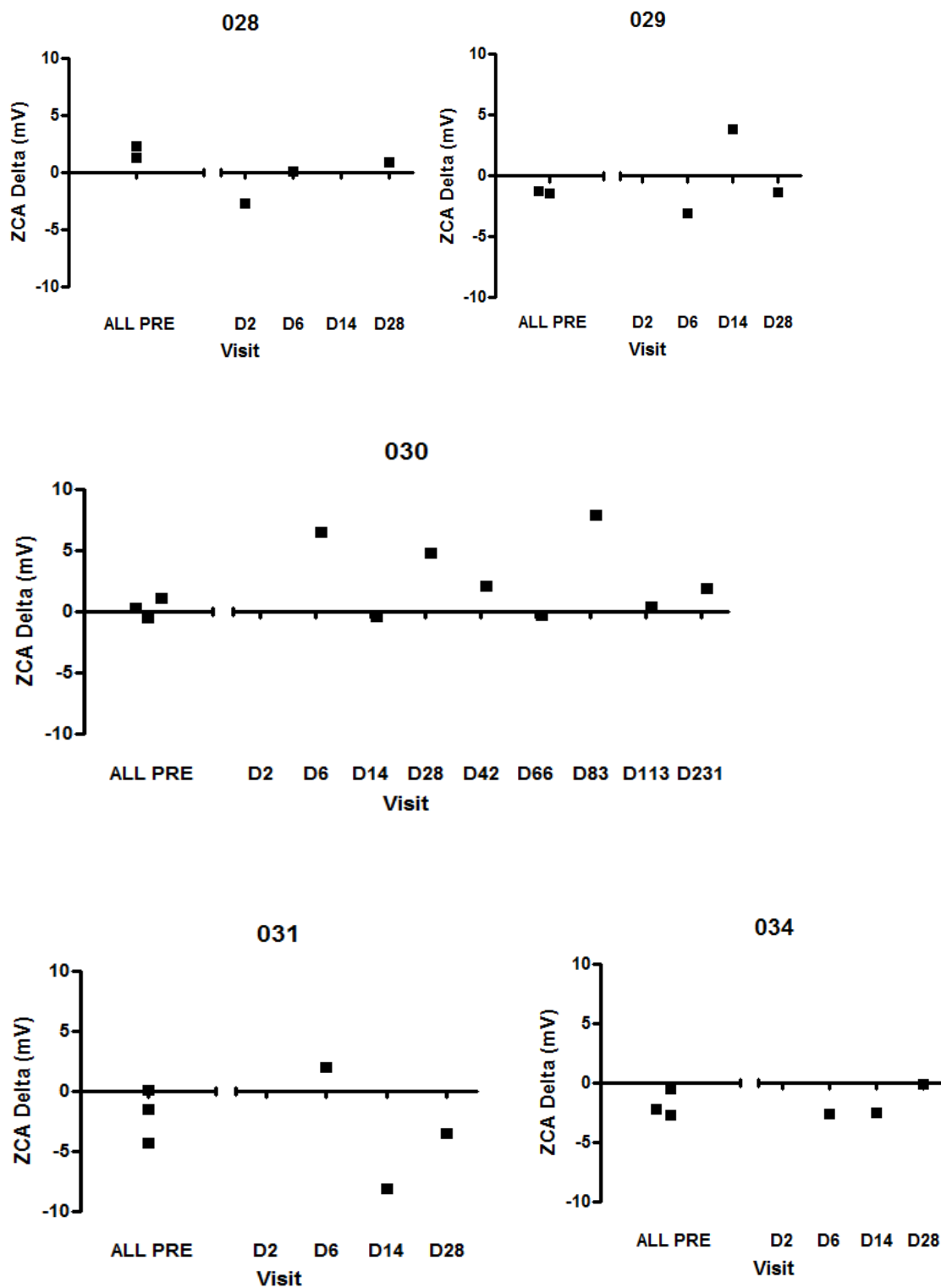
Figure 5.4. Paired results for Basal PD and response to Ringers amiloride perfusion. Results are for comparison of mean pre- with day 2,6,14 and 28 values. Figure 5.4A-D are for basal PD; and 5.4E-H for response to Ringers amiloride ('Delta RA'). No differences at any time point between pre- and post dosing were significant for basal results. There was a significantly greater depolarisation to RA (in an opposite direction to that of efficacy) between pre- and day 6 (Wilcoxon matched pairs test $p=0.02$), no differences at any other time point.

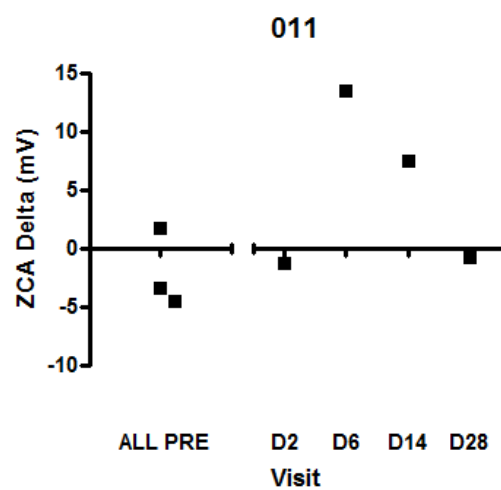
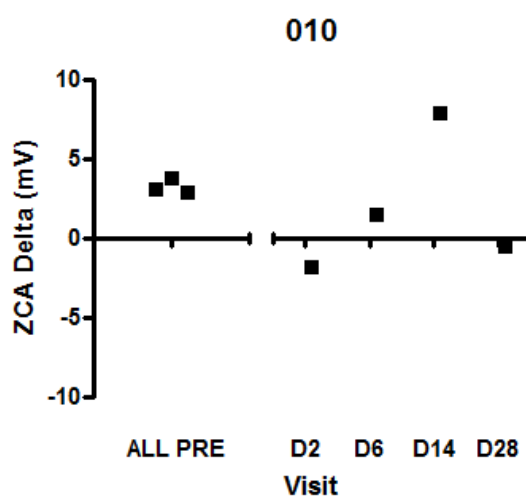
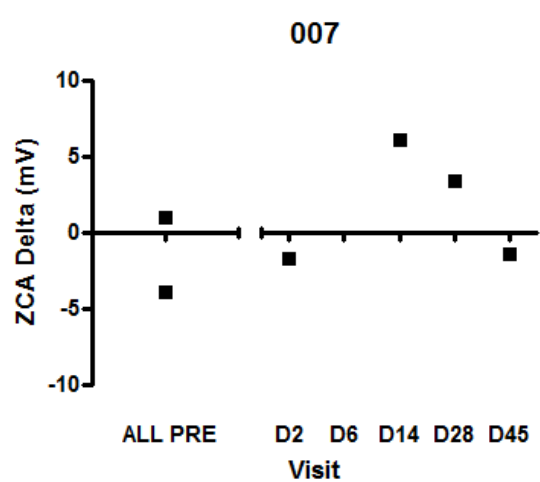
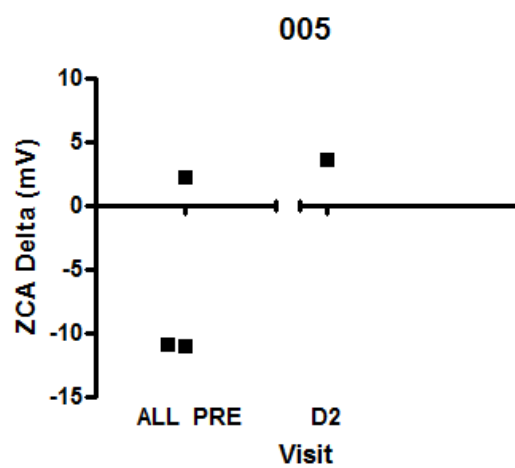
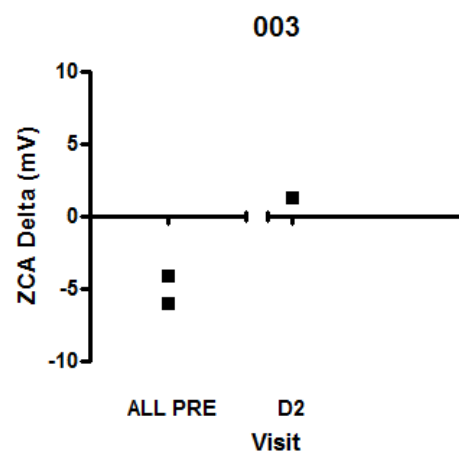
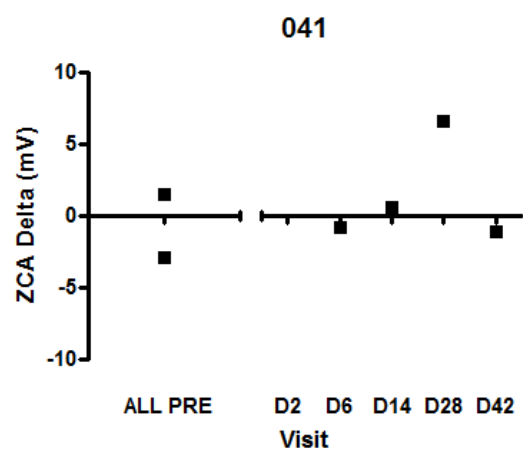
5.2.4.4 Chloride response

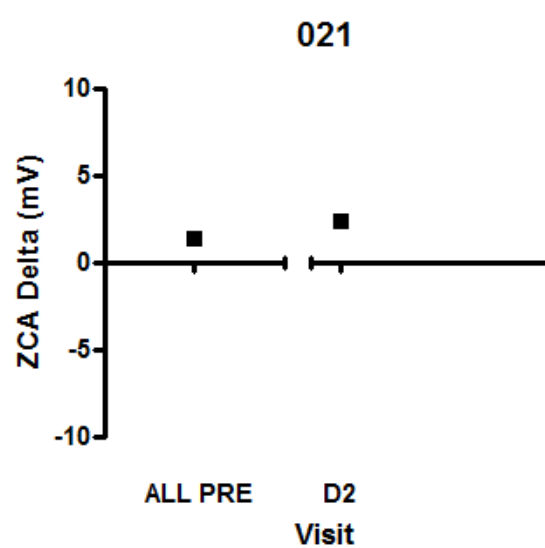
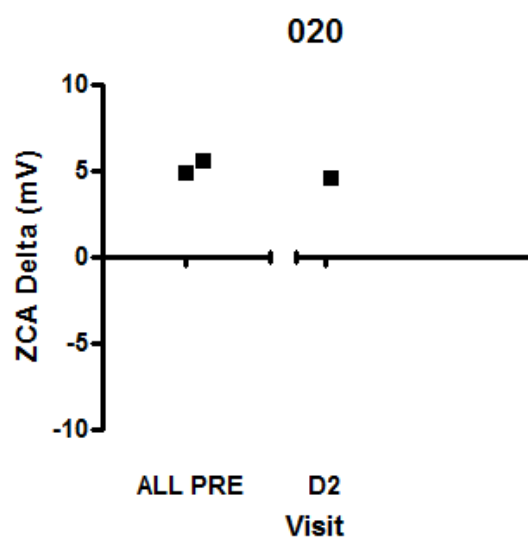
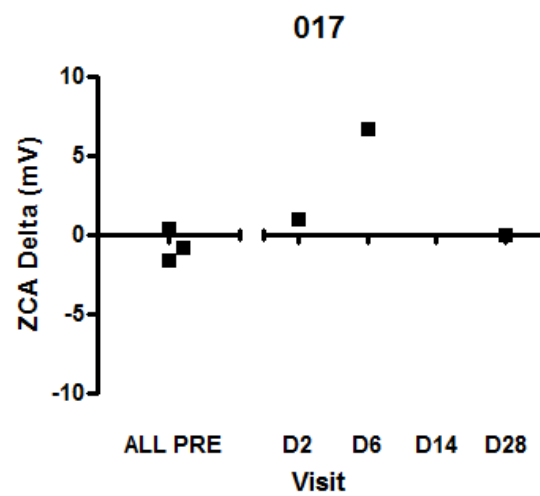
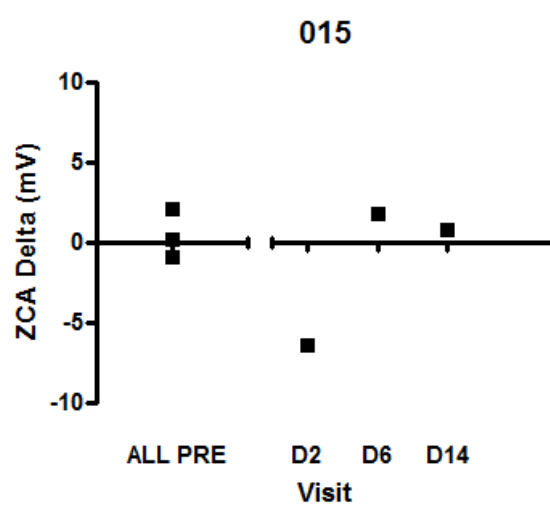
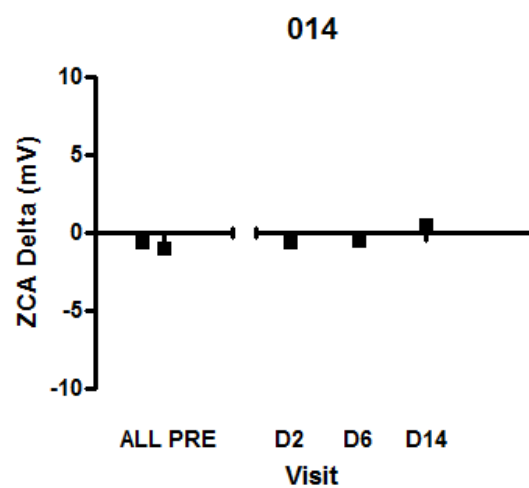
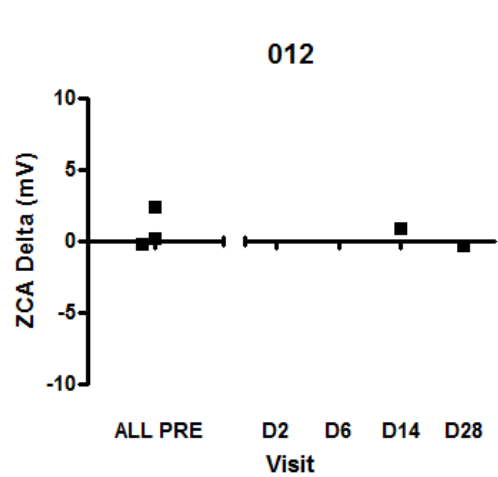
The change in PD ('delta') with zero chloride solution containing amiloride (ZCA) perfusion (to assess intrinsic chloride secretion) and for the sum of the ZCA and ZCAI sections (to assess total chloride transport) are presented according to timing of post dose measurement. An overall summary of results (according to the three definitions of 'response' in Chapter 2, section 2.5.3.6) can be found at the end of this section.

5.2.4.5 Intrinsic chloride secretion (ZCA delta)

The individual results for each patient (including each pre-dose and all post dose results plotted separately) are presented in Figures below (Figure 5.5).







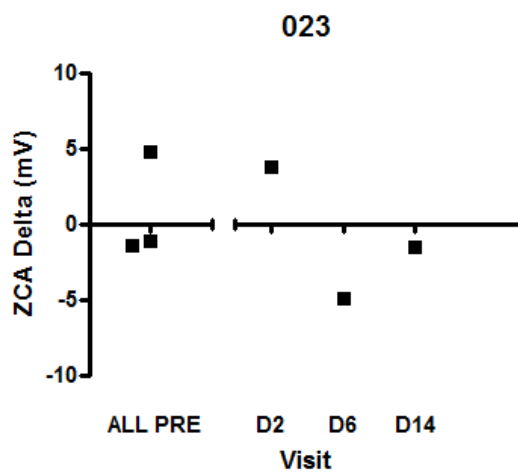
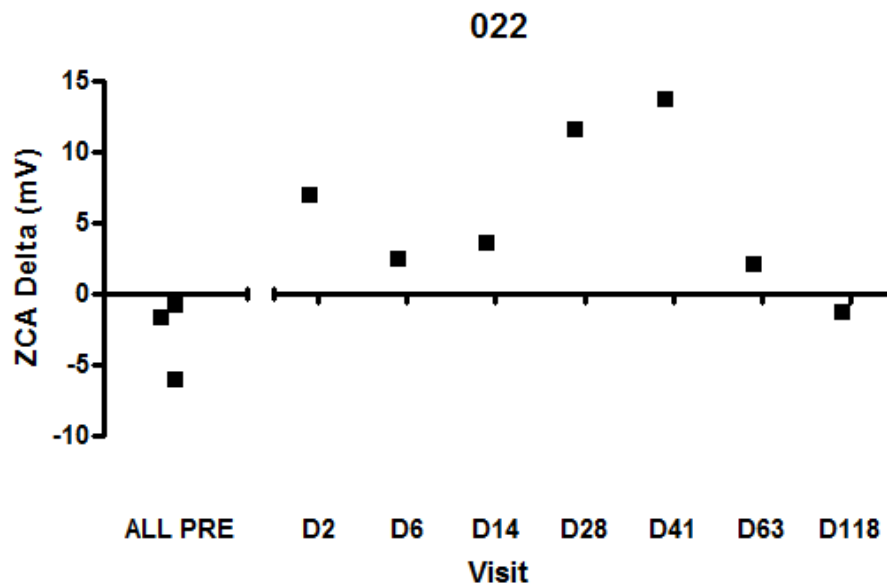


Figure 5.5. Intrinsic chloride secretion at all pre- and post dosing tests for the n=19 patients receiving a 2ml nasal dose. Patients identified according to study number as before. ZCA: Zero chloride with amiloride; Dx: day 'x' post nasal gene therapy dose.

The change in PD following ZCA solution perfusion is summarised in Figure 5.6 for the overall cohort. This is presented for all valid measurements taken at day 2, 6, 14 and 28 post dose in comparison to mean pre-dosing values.

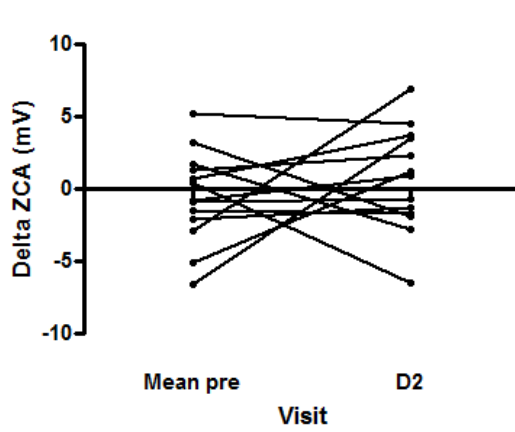


Figure 5.6A Delta ZCA to D2

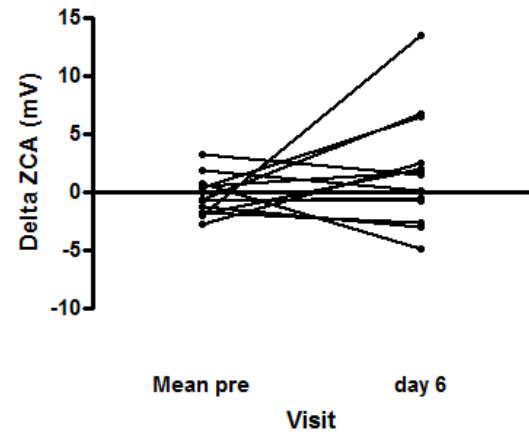


Figure 5.6B Delta ZCA to D6

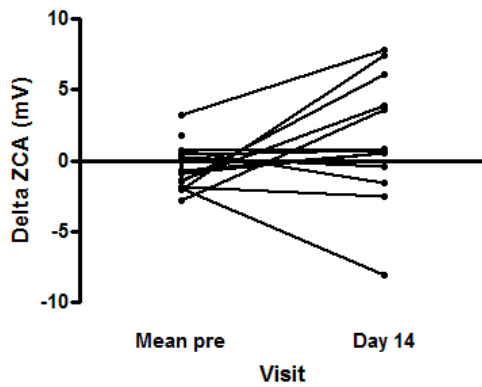


Figure 5.6C Delta ZCA to D14

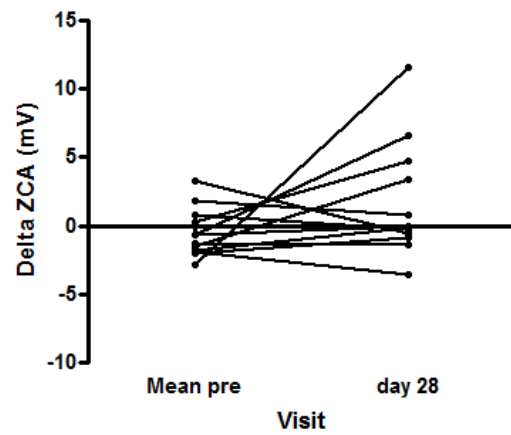
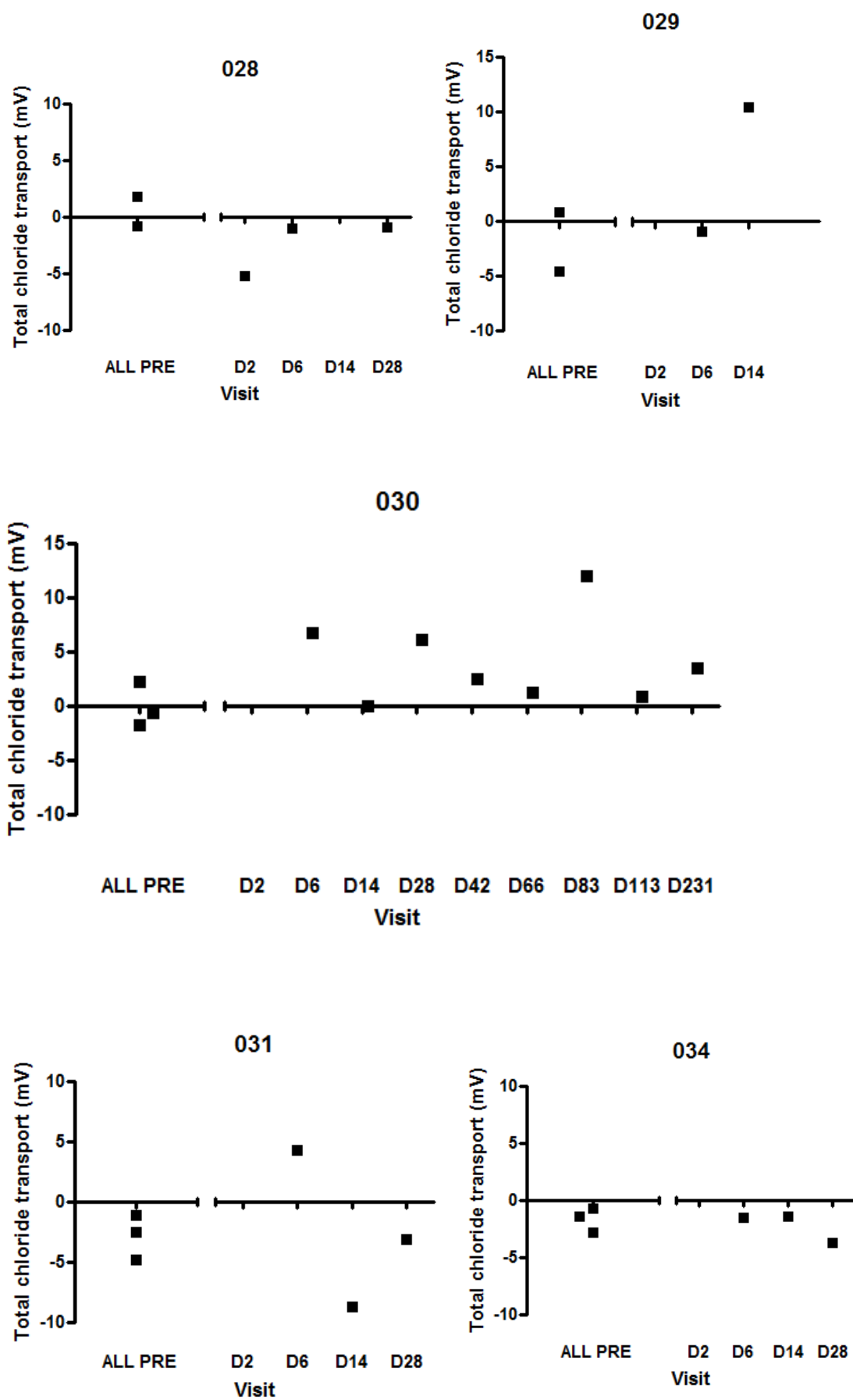


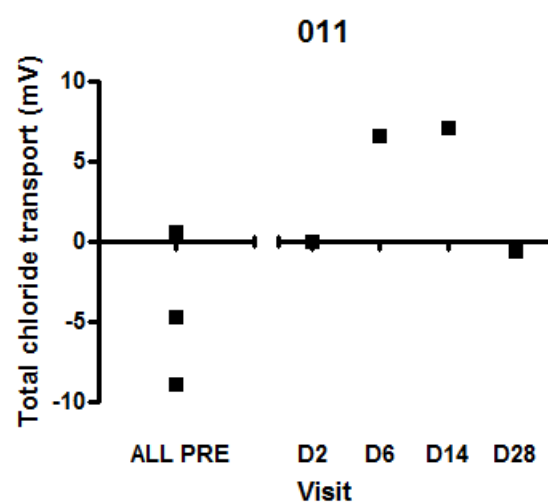
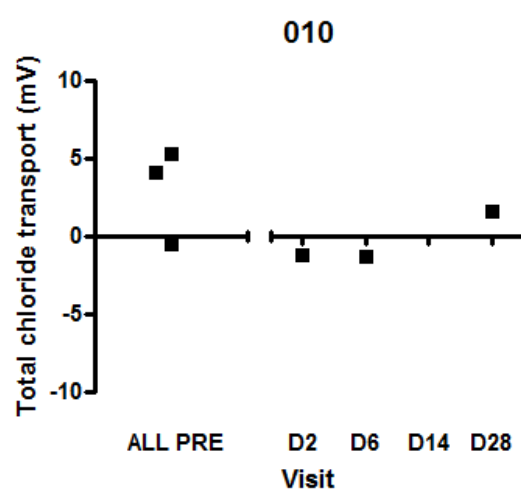
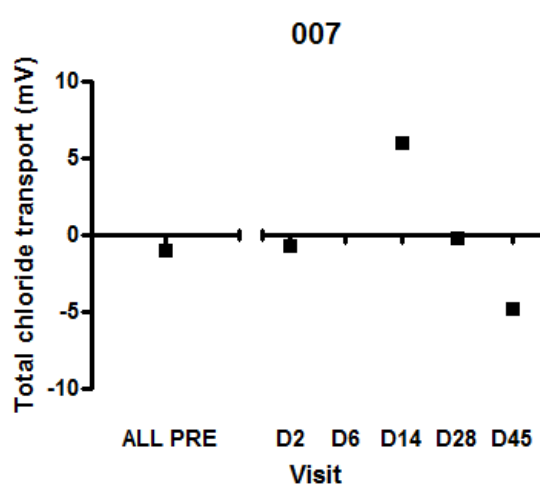
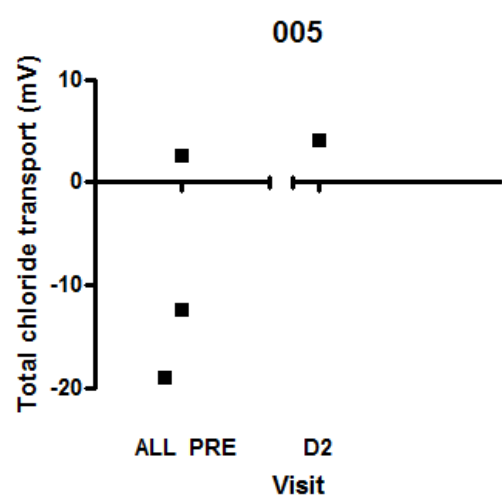
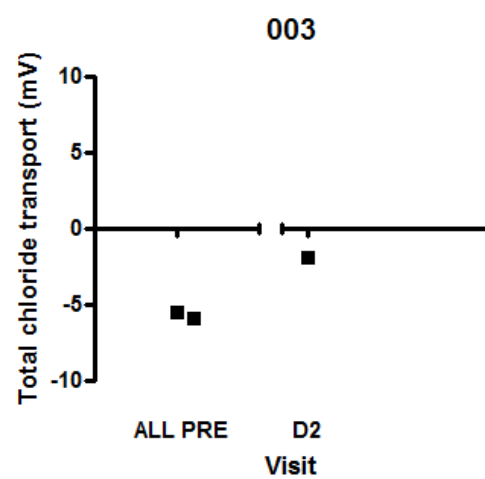
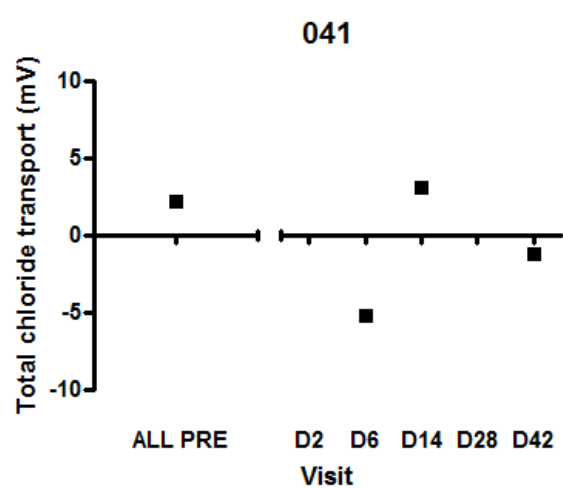
Figure 5.6D Delta ZCA to D28

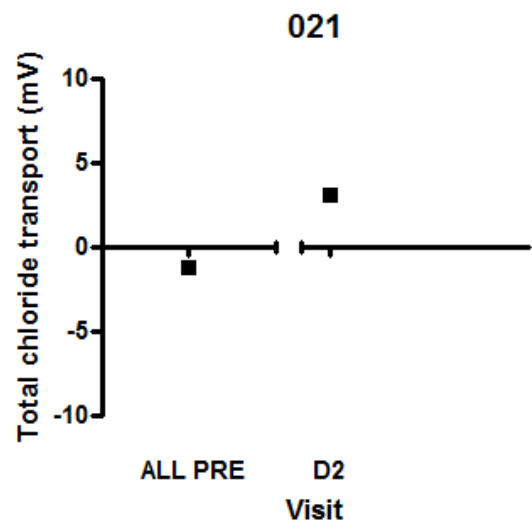
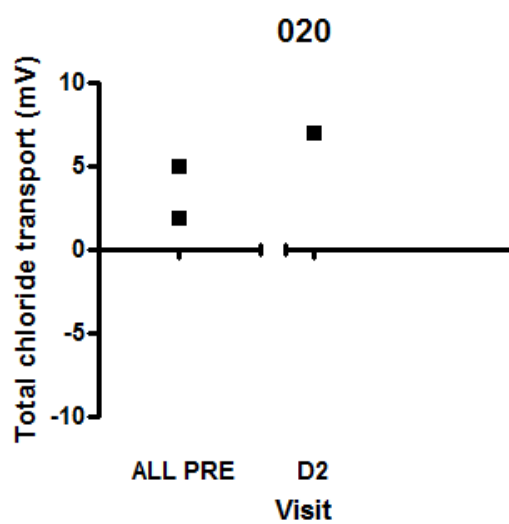
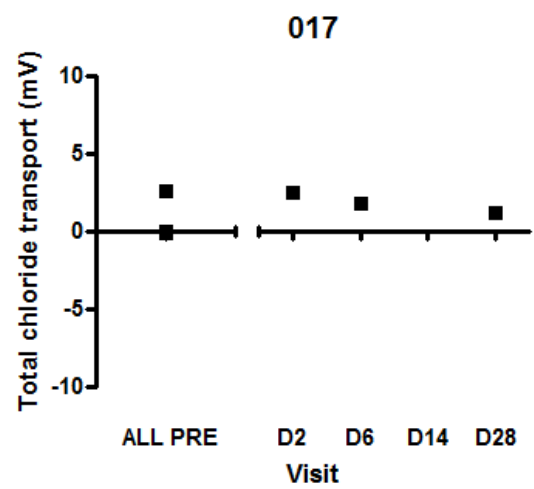
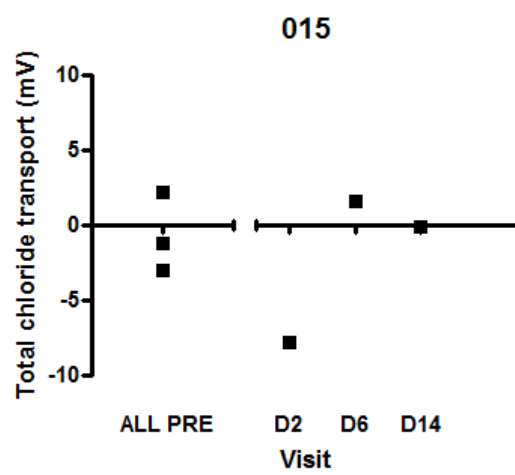
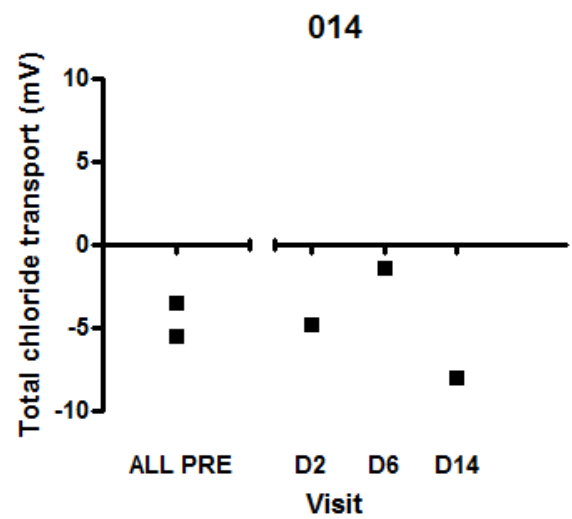
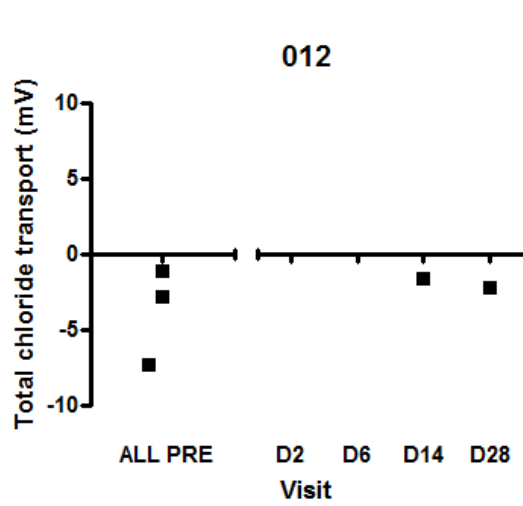
Figure 5.6. Intrinsic chloride secretion at post dose for patients with tests performed at day 2,6,14 or 28 (Not all patients had tests performed at each time point). Figures represent the different post dose time points as shown.

5.2.4.6 Total chloride transport (ZCA + ZCAI deltas)

The individual patient results for total chloride transport are shown in Figure 5.7 according to date of nasal PD. As per format for ZCA response, all pre- dosing results are plotted to demonstrate variability in measurement pre- intervention.







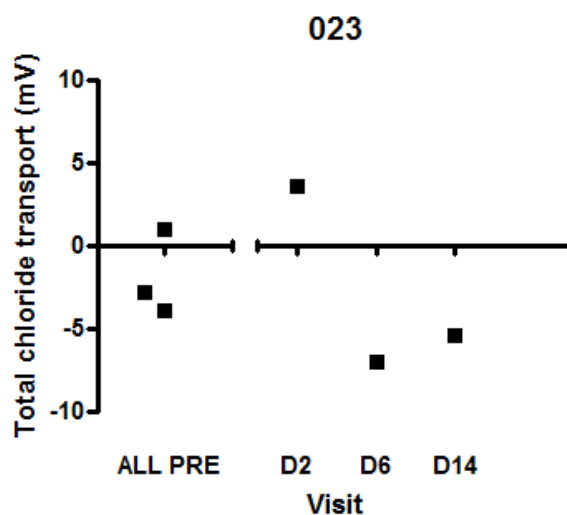
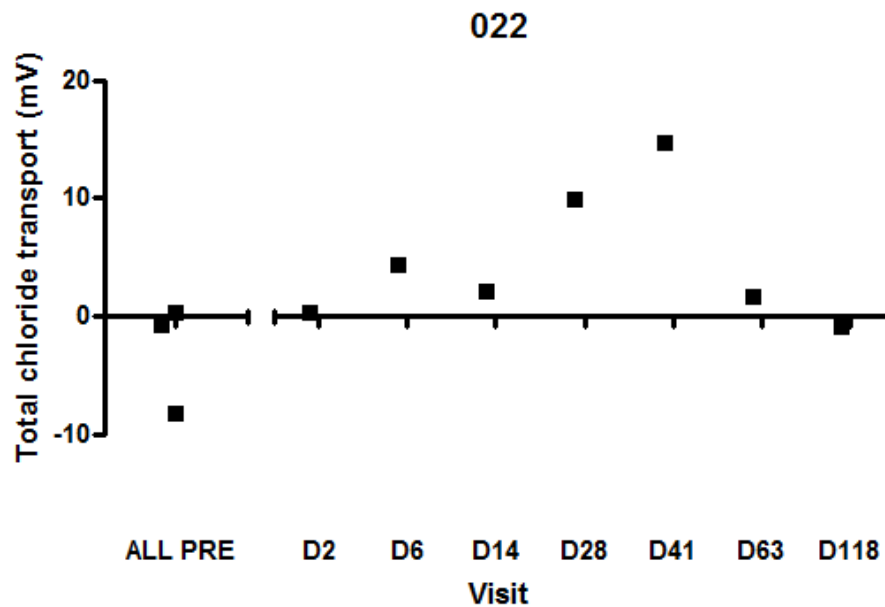


Figure 5.7. Total chloride secretion at all pre- and post dosing tests for the n=19 patients receiving a 2ml nasal dose. Patients identified according to study number as before. Total chloride secretion is the sum of deltas for zero chloride with amiloride and zero chloride with amiloride and isoprenaline response. Dx: day 'x' post nasal gene therapy dose.

To allow comparisons in overall group responses according to post dose time points, the results for total chloride transport are displayed in Figure 5.8, with the same format as applied to the presentation of intrinsic chloride secretion results above.

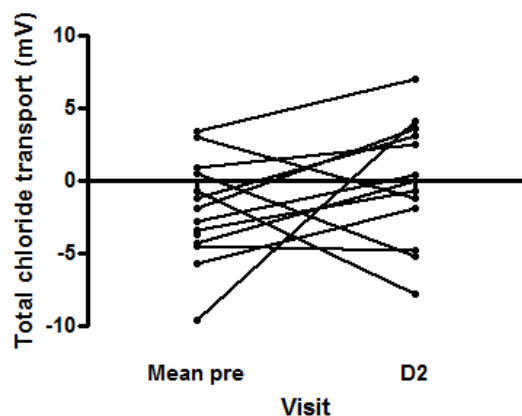


Figure 5.8A Total Cl⁻ transport to D2.

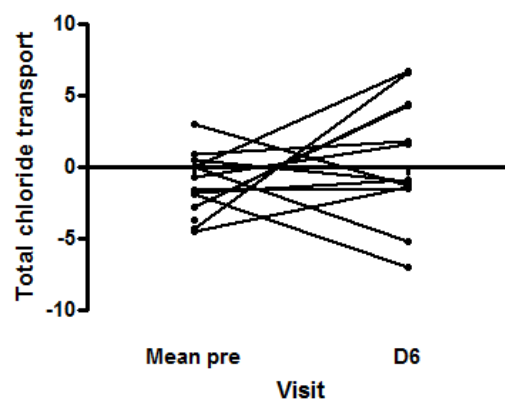


Figure 5.8B Total Cl⁻ transport to D6.

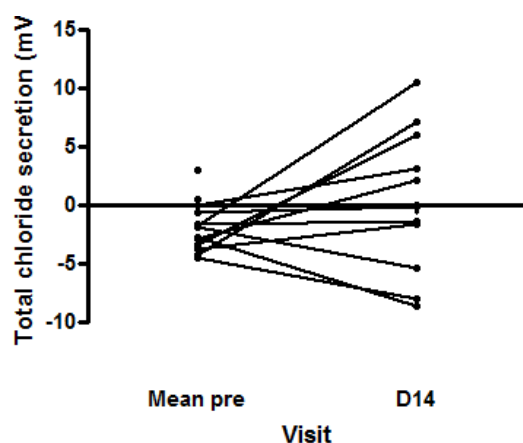


Figure 5.8C. Total Cl⁻ transport to D14.

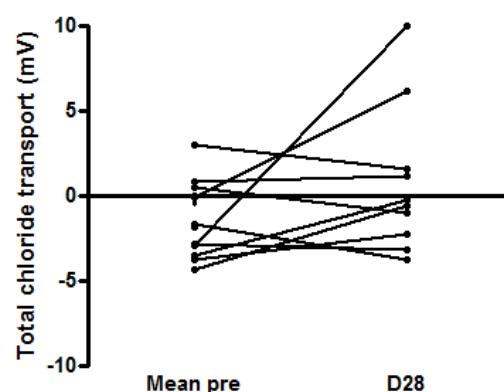


Figure 5.8D. Total Cl⁻ transport to D28.

Figure 5.8. Total chloride secretion at post dose for patients with tests performed at day 2,6,14 or 28 (Not all patients had tests performed at each time point). Figures represent the different post dose time points as shown.

5.2.4.7 Chloride secretion 'responses' according to definition

A summary of the proportion of patients meeting the criteria for a chloride secretion 'response' on a post dose nasal PD according to the definitions described in Chapter 2, section 2.5.3.6 is shown in Tables 5.5-5.8 according to the date of post dose test.

Day 2

	Method 1: Greater than any pre-dose test	Method 2: Change between mean pre- and post-dose test > 5mV	Method 3: Absolute chloride secretion > -5mV
Intrinsic chloride secretion (ZCA delta)	2/13 (15.4%)	3/13 (23.1%)	1/13 (7.7%)
Stimulated chloride secretion (ZCAI delta)	5/14 (35.7%)	1/14 (7.1%)	0/14 (0.0%)
Total chloride secretion (ZCA+ZCAI deltas)	4/13 (30.8%)	2/13 (15.4%)	1/13 (7.7%)

Table 5.5. Proportion of patients with a chloride ‘response’ according to each definition and section(s) of nasal PD test at day 2 post dose. Results are displayed for number of patients with ‘response’ according to definition of interest, and total number of patients with valid PD result for the section of PD of interest.

Day 6

	Method 1: Greater than any pre-dose test	Method 2: Change between mean pre- and post-dose test > 5mV	Method 3: Absolute chloride secretion > -5mV
Intrinsic chloride secretion (ZCA delta)	4/13 (30.8%)	4/13 (30.8%)	3/13 (23.1%)
Stimulated chloride secretion (ZCAI delta)	2/13 (15.4%)	0/13 (0.0%)	0/13 (0.0%)
Total chloride secretion (ZCA+ZCAI deltas)	5/13 (38.5%)	4/13 (30.8%)	2/13 (15.4%)

Table 5.6. Proportion of patients with a chloride ‘response’ according to each definition and section of nasal PD test at day 6 post dose. Results are presented as described for day 2 tests.

Day 14

	Method 1: Greater than any pre-dose test	Method 2: Change between mean pre- and post-dose test > 5mV	Method 3: Absolute chloride secretion > -5mV
Intrinsic chloride secretion (ZCA delta)	4/13 (30.8%)	4/13 (30.8%)	3/13 (23.1%)
Stimulated chloride secretion (ZCAI delta)	3/12 (25.0%)	1/12 (8.3%)	1/12 (8.3%)
Total chloride secretion (ZCA+ZCAI deltas)	3/12 (25.0%)	4/12 (33.3%)	3/12 (25.0%)

Table 5.7. Proportion of patients with a chloride ‘response’ according to each definition and section of nasal PD test at day 14 post dose. Results are presented as described for day 2 tests.

Day 28

	Method 1: Greater than any pre-dose test	Method 2: Change between mean pre- and post-dose test > 5mV	Method 3: Absolute chloride secretion > -5mV
Intrinsic chloride secretion (ZCA delta)	4/12 (25.0%)	2/12 (16.7%)	0/12 (0.0%)
Stimulated chloride secretion (ZCAI delta)	1/10 (10.0%)	0/10 (0.0%)	0/10 (0.0%)
Total chloride secretion (ZCA+ZCAI deltas)	2/10 (20.0%)	2/10 (20.0%)	2/10 (20.0%)

Table 5.8. Proportion of patients with a chloride ‘response’ according to each definition and section of nasal PD test at day 28 post dose. Results are presented as described for day 2 tests.

5.2.4.7.1 Mean pre-dose values vs. post-dose results in patients with tests on Days 6, 14 and 28

Ten patients had nasal PDs performed pre-dose, and on D6, D14 and D28 post dose. Figure 5.9 shows the mean (SD) results for intrinsic and total chloride secretion at mean pre-, day 6, 14 and 28 tests.

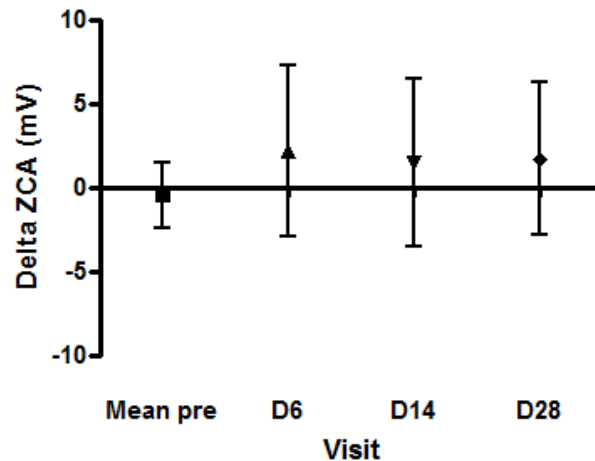


Figure 5.9A

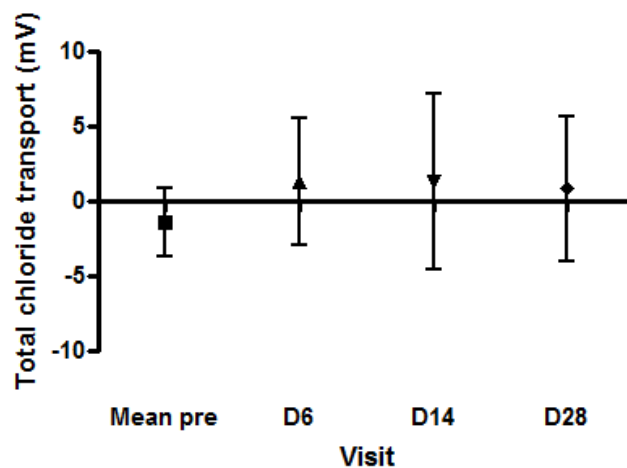


Figure 5.9B

Figure 5.9. Chloride transport at post dosing time points for all patients with tests at day 2, 6, 14 and 28 post dose. Figure 5.9A summarises intrinsic chloride secretion (ZCA delta) at mean pre-dose and day 2,6,14 and 28 post dose. Figure 5.9B summarises total chloride transport at mean pre-dose and day 2,6,14 and 28 post dose.

Because the pre-dose values for paired comparisons were means and because the time point of any changes seen post dose were not universal, the mean post dose values were also calculated to reduce some of the variability seen between tests within-subject. The mean pre- intrinsic (ZCA delta) and also total chloride secretion were compared to the mean of these responses for all patients with nasal PDs performed at day 6,14 and 28 post dose (data not shown). There was no significant difference between mean pre- results and mean [day 6,14 and 28] results for either intrinsic or total chloride secretion (paired t-test $p=0.11$ for both analyses).

5.2.4.7.2 Mean pre- to post-treatment values for individual patients

The definition used determined the number of patients with a chloride 'response'. At early time points (day 2 post dosing), the proportion of subjects with a 'response' was low according to the definition of response greater than any pre-dose test. Only 1SD003 and 1SD022 had a chloride secretion convincingly $>1\text{mV}$ outside pre-dosing values. Patient 1SD003 was in an early cohort of the 2ml nasal/20ml lung dose and therefore did not have PD measurements made at time points beyond day 2. In contrast, later patients did not have measurements performed at day 2 with the earliest post dose measurement at day 6.

A pattern was observed in individual patients (e.g. 1SD007) where a trace considered a 'response' (D14) was followed by staged return to 'non-response' (and therefore pre-dose) levels. In other patients, a different pattern was observed. Where a 'response' had been observed and a 'non-response' followed, further 'response' results were obtained at subsequent testing dates (see patient 1SD030 for example).

The greatest post-dose intrinsic chloride hyperpolarisations were 13.4mV (patient 1SD011, D6), 11.5mV and 13.7mV (from patient 1SD022, tests performed on day 28 and 41 respectively). The greatest post-dose total chloride hyperpolarisations were 10.5mV (patient 1SD029, day 14), 11.9mV (patient 1SD030 day 83) and 10.0mV and 14.8mV (from patient 1SD022, again tests performed on day 28 and 41 respectively).

For intrinsic chloride secretion (ZCA delta), 5/19 (26.3%) patients had at least one post dose nasal PD categorised as 'response' according to all three definitions and a further 4 patients met 'response' criteria according to two

definitions (and therefore overall 9/19 (47.4%) met 'response' according to $\geq 2/3$ definitions)(Table 5.9).

Patient	Intrinsic chloride secretion at any post dose test		
	Method 1: Greater than any pre-dose test	Method 2: Change between mean pre- and post- dose test > 5mV	Method 3: Absolute ZCA chloride secretion > -5mV
028	N	N	N
029	Y	Y	N
030	Y	Y	Y
031	N*	N	N
034	N*	N	N
041	Y	Y	N
003	Y	Y	N
005	N*	Y	N
007	Y	Y	Y
010	Y	N	Y
011	Y	Y	Y
012	N	N	N
014	N*	N	N
015	N	N	N
017	Y	Y	Y
020	N	N	N
021	N*	N	N
022	Y	Y	Y
023	N	N	N
Total	Y: 9/19 (47.4%) N: 10/19 (52.6%)	Y: 9/19 (47.4%) N: 10/19 (52.6%)	Y: 6/19 (31.6%) N: 13/19 (68.4%)

Table 5.9. Intrinsic chloride secretion 'response' post dose at any time point according to definition used. Y=meets criteria for response, N=does not meet criteria. *Although greater than pre-dose test, minimal difference and therefore considered non-response.

Similarly for total chloride transport (ZCA + ZCAI delta), 5/19 (26.3%) patients had at least one post dose nasal PD categorised as 'response' according to all three definitions. Four of these were the same patients meeting all 3 definitions for intrinsic chloride response (007, 011, 017, 022 and 030). A further 3 patients met 'response' criteria for total chloride transport according to two definitions (and therefore overall 8/19 (42.1%) met 'response' according to $\geq 2/3$ definitions)(Table 5.10).

Patient	Total chloride secretion at <u>any</u> post dose test		
	Method 1: Greater than any pre-dose test	Method 2: Change between mean pre- and post- dose test > 5mV	Method 3: Absolute total chloride secretion > -5mV
028	N	N	N
029	Y	Y	Y
030	Y	Y	Y
031	Y	Y	N
034	N	N	N
041	N*	N	N
003	Y	N	N
005	N*	Y	N
007	Y	Y	Y
010	N	N	N
011	Y	Y	Y
012	N	N	N
014	Y	N	N
015	N	N	N
017	N	N	N
020	Y	N	Y
021	Y	N	N
022	Y	Y	Y
023	Y	Y	N
Total	Y: 11/19 (57.9%) N: 8/19 (42.1%)	Y: 8/19 (42.1%) N: 11/19 (57.9%)	Y: 6/19 (31.6%) N: 13/19 (68.4%)

Table 5.10. Total chloride secretion ‘response’ post dose at any time point according to definition used. Y=meets criteria for response, N=does not meet criteria. *Although greater than pre-dose test, minimal difference and therefore considered non-response.

PD responses at extended time points were investigated in some subjects with a trace seen at the day 28 visit which, on visual inspection, was consistent with a ‘response’, and who were willing to attend further visits for testing. This applied to four patients. Two of these patients had evidence of sustained changes in chloride secretion beyond the original series of nasal PD time points to day 28, as shown previously in the individual graphs for intrinsic and total chloride secretion (patients 1SD022 and 030).

5.2.5 Actual nasal dose (absolute volume) delivered and evidence of CFTR function

The actual volume of nasal gene therapy delivered following the administration of the specified number of nasal spray actuations is described in Chapter 4, section 4.3. There was no correlation between the actual volume of gene therapy delivered and the likelihood of a ‘response’ (Table 5.11).

	Volume administered (ml) mean (SD)		Comparison
	Post dose nasal PD response*	No nasal PD response	
Intrinsic chloride secretion	2.0 (0.4)	1.8 (0.2)	p=0.45
Total chloride secretion	2.0 (0.2)	1.8 (0.4)	p=0.12

Table 5.11. Actual volume of nasal gene therapy delivered (2ml dose cohort) and nasal PD outcome for intrinsic and total chloride secretion. *criteria met according to 2 or more definitions at any post dose test. Unpaired t-tests were used to compare differences in response according to actual volume administered.

5.2.6 Comparison with other CFTR outcomes

5.2.6.1 Nasal PD and molecular results

As described in Chapter 2, section 2.7, patients in the nasal dosing cohort had a nasal brushing at a single time point post dose for vector specific *CFTR* mRNA. No patients receiving a 1ml nasal dose had a positive mRNA result. Three patients in the main 2ml cohort had detectable mRNA in their post dose test (although positive, the values were not quantifiable)(Table 5.12). There was no association between the nasal PD response and the detection of *CFTR* mRNA.

Patient	Intrinsic chloride 'response' on nasal PD?*	Total chloride secretion 'response' on nasal PD?*	Vector specific <i>CFTR</i> mRNA
028	N	N	Positive (D28)
029	Y	Y	Not detected
030	Y	Y	Not detected**
031	N	Y	Not detected
034	N	N	Not detected
041	Y	N	Not detected
003	Y	N	Not detected
005	N	N	Not detected
007	Y	Y	Not detected
010	Y	N	Not done
011	Y	Y	Positive (D14)
012	N	N	Not detected
014	N	N	Not detected
015	N	N	Not detected
017	Y	N	Not detected
020	N	Y	Not detected
021	N	N	Positive (D14)
022	Y	Y	Not detected
023	N	Y	Not detected

Table 5.12. Summary of nasal PD chloride secretion outcome and molecular evidence of vector specific *CFTR* on post 2ml dose nasal brushing. *criteria met according to 2 or more definitions at any post dose test. Not detected: Sample taken but no vector specific mRNA detected. **Two post dose brushing samples taken, one at day 28 and one at an extended time point > day 28, both were negative for mRNA. Not done: no post dose sample (patient refused nasal brushing).

5.2.6.2 Nasal PD and lower airway PD

Ten patients had pre- and post-dose bronchoscopies during which measurements of lower airway PD were made. Nasal PD tests were carried out on the same day. Associations between upper and lower airway PD results are described in Chapter 6.

5.2.7 Characteristics of patients with chloride response on post dose nasal PD

The genotypes of all 22 patients receiving a nasal dose of gene therapy are described in Chapter 4, section 4.2.2; all patients receiving the 2ml nasal dose had at least one p.Phe508del mutation. A summary of patient characteristics

including genotype, gender, age and lung function at recruitment is shown in Table 5.13. There was no difference in genotype distribution, gender, age or FEV₁% predicted at recruitment between patients with or without evidence of an intrinsic nasal PD ‘response’ on one or more post dose tests. There was a significant difference in gender distribution according to the presence of a total chloride ‘response’ but the small number of females in the cohort (n=5; 26%) this result should be interpreted with caution.

	Intrinsic chloride ‘response’ on nasal PD?*		Total chloride secretion ‘response’ on nasal PD?*	
	Yes	No	Yes	No
Genotype				
p.Phe508del/p.Phe508del	5	6	3	8
p.Phe508del/other	2	2	3	1
p.Phe508del/unknown	2	2	2	2
Gender				
Male	7	7	8	6
Female	2	3	0	5
Age (years), median (IQR)	31.0 (25.6-52.7)	25.9 (21.8-28.5)	35.4 (25.2-54.7)	26.8 (23.0-28.6)
FEV ₁ % predicted at Initial visit, median (IQR)	69.9 (63.7-86.9)	73.8 (66.1-85.3)	68.4 (63.9-82.1)	74.7 (66.2-85.9)

Table 5.13. Patient characteristics according to evidence of chloride ‘response’ on nasal PD (2ml dose cohort). *criteria met according to 2 or more definitions at any post dose test. Comparisons made between ‘response’ and ‘no response’ with Mann Whitney and Fishers exact tests; all were non-significant (p>0.05) for the presence of an intrinsic chloride response. For total chloride secretion, there was a significant difference between evidence of chloride response vs. non-response for gender (Fisher’s exact test p=0.04) but not for age or FEV₁% predicted (Mann Whitney p=0.3 and p=0.6 respectively).

5.3 Novel approach to nasal PD

5.3.1 Measurements in the non-CF and CF nose

The results described in section 5.2 were all performed with the LR4 voltmeter and nasal PD set up (Chapter 2, section 2.5.1-2.5.2). A novel approach to nasal PD testing was investigated using the LR5 device (Chapter 2, section 2.5.3). The PD measurements obtained using the LR4 were identical to readings obtained from the LR5 prior to clamp application. Application of a clamp was able to influence the measured PD, with an effect size dependent on the magnitude of

clamp applied. There was no relationship between clamp duration and the rate of return (number of data values collected at 33 Hz before return to pre-clamp values) to basal PD. The change in PD (post standard nasal PD testing protocol including solution perfusion) in response to different clamp settings is shown in Figure 5.10.

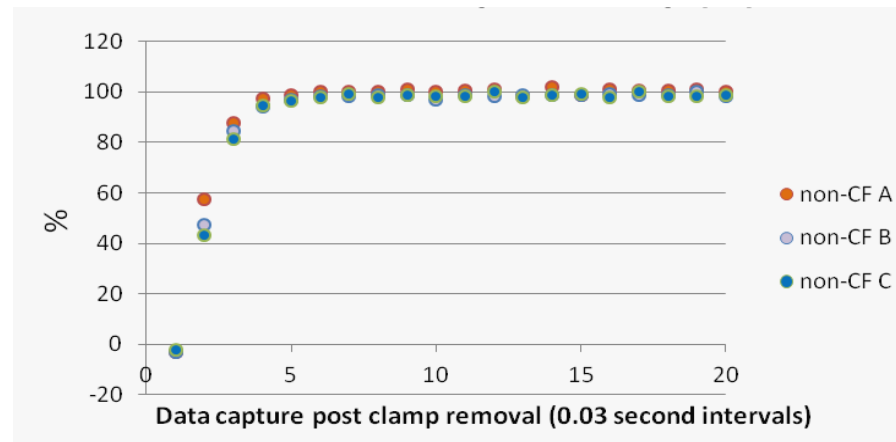


Figure 5.10. Recovery to baseline PD following application of maximal clamp setting. Triplicate measurements in n=2 subjects (one subject with 2 tests). Data was captured every 0.03 seconds, and it can be seen that return to baseline (pre-clamp) PD was achieved within 0.12 seconds.

The recovery time post clamp removal was almost instantaneous (within the 33Hz detection capabilities) irrespective of filter setting used for CF and non-CF subjects. There was no difference in rate of return to pre-clamp PD post clamp removal between non-CF and CF. In non-CF subjects, there was no difference between traces obtained either in isolation or immediately following a standard LR4 nasal PD (i.e. post perfusion steps).

One of the CF patients included in the CF group had a significant chloride response on ‘standard’ nasal PD testing. Therefore only one nasal PD (LR4 and LR5) was carried out on this subject as he was subsequently excluded from the nasal dosing cohort. There was substantial overlap between CF and non-CF subjects in recovery of PD post clamp removal, irrespective of whether the traces were obtained immediately following or in isolation of a standard PD. Regardless of filter setting, return to baseline PD was almost instantaneous within the detection capabilities of the hardware. All results for patients with CF completely overlapped those from non-CF subjects.

The change in mV expressed as a % of baseline PD during clamp application was no different pre- and post- administration of a single dose of CF gene therapy (a representative result from patient 1SD028 is shown in Figure 5.11). Following release of the clamp return to baseline was within the first 2-3 data capture points (within 0.1 seconds) in all tests performed.

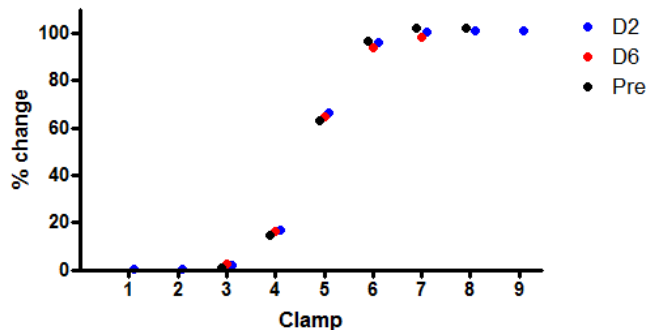


Figure 5.11. Change in baseline PD in response to clamp setting. Clamp 9 is the maximal load. Increasing load changed the measured PD consistently between tests irrespective of relationship to gene therapy dose. Pre = pre dose; D2 and D6 refer to day 2 and day 6 post dose respectively.

5.3.2 Further experimental work

The results confirmed that the application of an electrical ‘load’ could influence the PD measured across the respiratory epithelium, but the characteristics were not different between CF and non-CF individuals. It was hypothesised by the author that changes in PD post clamp application could reflect high impedance of the entire epithelial cell and test equipment ‘circuit’ under test, in particular, that the Foley catheter containing the nasal electrode was a potential source of high impedance.

5.3.2.1 Electrode type

In order to exclude the high impedance components of the electrode / electrode cream / Foley catheter combination at the exploring electrode, an alternative electrode was sourced. A nasopharyngeal electrode (DW020, Digitimer, UK) CE marked for use *in vivo* during electrode encephalogram (EEG) testing met the criteria required: the electrode was comprised of a silver/silver chloride electrode small ball (not disk), on the end of the semi-rigid coated wire which

could be inserted directly into the nose and negate the requirement for a Foley catheter. The electrode comprises the recording ball 'tip' (3mm diameter) and 1.5m silicone wire (Figure 5.12). A disk electrode alone was unsuitable due to the diameter and discomfort on insertion. As no perfusion steps are carried out during tests with the LR5, it was not necessary to have a channel for perfusion which is the purpose of the Foley catheter during standard nasal PD testing.

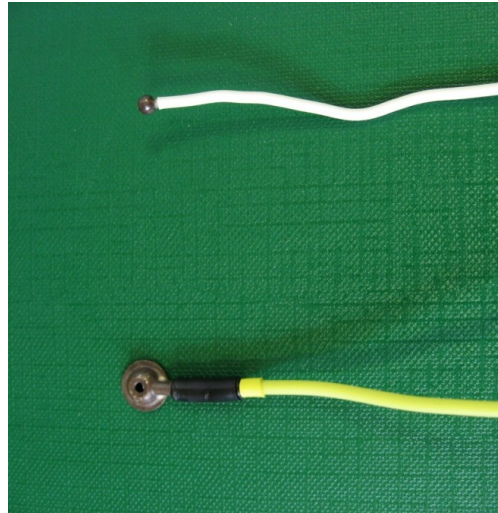


Figure 5.12. Silver/silver chloride ball electrode. This photograph shows the ball electrode (top) with the disk electrode used in our existing nasal PD testing set up (bottom) for comparison.

This electrode was compared with the standard nasal electrode set up (disk electrode in Foley catheter as described in Chapter 2, section 2.5) in a non-CF subject. The nasopharyngeal electrode was inserted along the floor of the nasal cavity to the site of most negative PD (3cm into nose). The ball tip of the electrode had been dipped in electrode cream prior to insertion to improve connectivity with the nasal epithelium. F1 and WB filtering of the measured PD were used, with maximal clamp (setting 9) and clamp time (10 seconds).

5.3.2.2 Results

No differences were observed between n=3 tests at each clamp setting (load) or signal filter setting performed with the disk electrode in Foley catheter vs. Ag/AgCl ball electrode for the single subject (author). The electrode caused discomfort whilst in situ and therefore further testing in other subjects was not carried out. Comparison with the original nasal electrode / catheter arrangement is shown in Figure 5.13. Overall, no difference was observed with

the use of the direct nasopharyngeal electrode in comparison to the standard electrode set up.

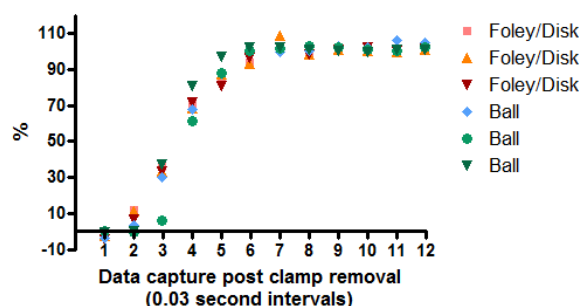


Figure 5.13. Comparison of electrode set up with LR5. Each data point represents the mean of 3 values, performed with the original electrode and Foley catheter and disk electrode set-up, and with the nasopharyngeal ball electrode. Each test was performed in triplicate as shown.

From the initial test results, none were obtained to support the hypothesis that patients with CF would have a slower return to baseline PD than non-CF subjects, and no change was seen in the limited testing points post nasal gene therapy. However of note, no post dosing tests were performed in patients with the LR5 at time points when a change in chloride transport was observed with standard PD testing, and results obtained were limited to a small number of subjects. Despite these limitations, the absence of any CF/ non-CF difference and an additional time factor added to the Pilot nasal PD for patients as a result of testing with the LR5 meant that a decision was taken by the author not to pursue further investigation in the Pilot study.

5.4 Discussion

This chapter has described the nasal PD results performed in the CF gene therapy Pilot study. Results demonstrated that on an individual patient level, a single nasal dose of CF gene therapy (pGM169/GL67A) can result in detectable changes in chloride secretion on nasal PD testing with a perfusion protocol including Ringers, RA, ZCA and ZCAI solutions. Although some non-significant trends towards a ‘response’ on a group level were seen, the PD changes are not universal. In some patients, prolonged changes in PD in response to zero chloride

solutions were seen beyond initial expectations at trial onset. Yet in others, who were not obviously different in terms of genotype, disease severity (as assessed by FEV₁), gender and age, it did not change. This chapter has also described results from a preliminary investigation of a novel approach to PD testing with which the author had hoped to improve the quantitative ability of PD to detect change post gene therapy (section 5.3). It was felt by the author that on the basis of the initial results that further testing in the Pilot study was not justified, despite only performing post dose measurements in n=2 subjects. This is discussed in more detail below.

5.4.1 Feasibility of airway PD testing

When considering appropriate outcome measures for early phase clinical trials, it is important that patient, personnel and economic factors are taken into consideration. Although it is not painful, nasal PD measurement is not considered by most patients a particularly pleasant test. It involves sitting still without talking for 20-30 minutes and a slow constant dripping of perfusate from the nose during this time. Patients tolerated repeated measurements in the Pilot study very well. However, more frequent testing may not be as well tolerated.

With the (necessary and important) GCP modifications to the PD set up as described in Chapter 2, section 2.5.1-2.5.2, nasal PD changed from being a test with relatively low running costs to becoming extremely expensive. This was further complicated with repeated testing necessitated by uncertainty over the kinetics of gene expression in a ‘first in man’ study, the inherent within subject variability and absence of a uniform response (or pattern of response).

Studies such as the Pilot study may help better guide appropriate time lines for testing, but for ‘first in man’ studies protocol design can only be guided by the results of pre-clinical studies. Flexibility within the trial protocol is important, but nasal PD will always be a cross-sectional test of a dynamic process over a small area of tissue and at intermittent points in time.

For nasal PD to be adopted more widely as an outcome measure in the UK, appropriate training for investigators is important. The LR4 voltmeter set up used in the Pilot study was not CE marked and therefore to meet regulatory requirements, responsibility was transferred from the inventor to the RBH

clinical engineering department. For studies involving centres outside RBH this might not be acceptable to relevant regulatory authorities, and in the absence of CE approval, alternative hardware for PD testing might need to be sought (such as the PowerLab (AD Instruments, Colorado, US)), as recommended by the US Therapeutic Development Network Standard Operating Procedure (TDN SOP) (110)). Furthermore, there is an issue of technical experience and expertise. All nasal PDs in this study were performed either by the author or JD. Each test takes around 30 minutes to perform but the times involved in test preparation are significantly greater. Similarly, data analysis takes around 25 minutes per nasal PD test for a single analyser. Although likely to have a greater role to play in the clinic in the future, at present only a limited number of tests are performed outwith the trial setting. Opportunity for learning, and adoption of a standardised technique would be required in Europe, similar to that achieved in the US with their TDN approach and indeed standardisation process is taking place currently within the European CF Society Clinical Trials Network (94). However, it is important to bear in mind that increasing the number of sites and operators is likely to increase variability in results.

In the UK CFGT Consortium's current multidose trial, PD is being undertaken using the methodology described in Chapter 2, section 2.5.1-2.5.2 at a single (London) site (RBH) despite patients being recruited from both the London and Edinburgh sites, in an effort to avoid the variability which would be introduced by multi-site testing.

5.4.2 Strengths of design

In line with the published literature (94), we observed variation in pre-dosing tests at all stages of the test (section 5.2.4.1). However, by increasing the number of pre-dose tests from the original study reported by Alton et al (35)(the 'Lancet study') where only one pre-dose test was performed, there is better opportunity to be able to detect real change and understand the internal variability of PD within a subject over time. Use of the mean pre-treatment values was not always optimal in determining 'response'. This was because a post dose result could be 'greater' than mean pre-dose levels yet still be within the overall range of pre-dose values.

We undertook up to three pre-dosing nasal PD measurements in study participants and in the latter phase of the trial excluded patients with evidence of chloride secretion considered within the normal range from the nasal arm of the study. There are examples in the literature where nasal PD outcome post intervention is compared with only one pre-dosing measurement both for studies of CF gene therapy and small molecule treatments (35;36;96;97). Variability in nasal PD between pre- and post-dosing results may reflect *de novo* CFTR expression and resultant changes in chloride secretion, or just expected variability within that patient (irrespective of any intervention) hence the importance of being able to have as much information as reasonably possible on pre-dosing measurements. Variation in results due to experimental set up and personnel undertaking the test are of considerable importance. In the US, the development of the TDN SOP (110) for nasal PD measurement was in response to considerable variability seen between 4 different centres involved in a Phase I trial evaluating a potential *CFTR* corrector (8-cyclopentyl-1,3-dipropylxanthine [CPX]) (85) as reported by Ahrens et al (84). This trial failed to detect change in treated patients on the basis of nasal PD results. Rowe et al (109) have collated the results from the CPX trial reported by Ahrens et al and compared it to a different multicentre trial performed post development of the US TDN SOP (topical aminoglycosides, analysis performed involving n=5 sites using the TDN SOP)(123). Total chloride secretion was similar, yet the variance was greatly reduced post SOP (CPX mean total chloride secretion 1.9mV, SD 6.6; nasal aminoglycosides mean total chloride secretion 2.6mV, SD 2.97) (109).

The basal PD values observed in the Pilot study are similar to those reported by Middleton et al (75) and by other groups reporting measurements in CF subjects, recently collated by the European Cystic Fibrosis Society Clinical Trial Standardisation Committee (94).

The nasal PD data collection in the Pilot study was also strengthened by the nasal PDs being performed by only 2 operators and being a single site study. Therefore there were no additional variables such as equipment set up or PD solution composition. In addition, all data analysis was performed blinded to subject and chronology. A robust process of data analysis and high levels of inter-reporter agreement between tests (further improved post consensus review) as described in section 5.2.2 added to the quality of data. I considered

performing an inter-reporter agreement for whether GD or JD had performed the nasal PD within individual patients. However, there were several important confounding factors (such as the proportion of tests performed within subjects by each operator) which meant that I considered such an analysis was not valid.

In Pilot part B, we excluded patients from nasal dosing if they had evidence of a significant ($>5\text{mV}$) hyperpolarisation response to zero chloride solution perfusion. Inclusion of patients with evidence of significant hyperpolarisation can make interpretation of post intervention results challenging. Leonard et al have recently reported the results of a randomised placebo controlled crossover trial of miglustat (N-butyldeoxynojirimycin, Zavesca®) which used change in total chloride secretion on nasal PD as the primary outcome measure (124). There was no evidence of a treatment effect, however at baseline measurement half of the patients had evidence of significant chloride secretion, potentially influencing trial outcome. They recommended excluding patients with evidence of chloride secretion for future trials in combination with improving standardisation of data collection and analysis (124). Our practice to exclude such patients in the latter phase of the Pilot study is continuing for participation in the nasal cohort of the UKCFGTC multidose trial.

5.4.3 Limitations of nasal PD procedure and trial design

Although it must be emphasised that the trial was not intended to be powered for change in any PD parameter on a group level, the evaluation of PD response within a pre-determined protocol limited the duration of post-dose testing for patients in part A of the study. However, the results provide a useful framework for an evaluation of a single nasal dose of CF gene therapy product pGM169/GL67A.

With the benefit of hindsight, the Pilot trial protocol could have been designed to better allow the opportunity to capture any potential long duration changes. Time points for testing were not uniform between subjects in view of the iterative trial design. Trial design was driven mainly by time points for safety outcome measures, with time points for efficacy end points were based (appropriately) on best available evidence at the time of trial design (35). For the patients receiving a 1ml nasal dose and the first 3 patients receiving a 2ml

nasal dose, the last nasal PD was on day 2 post treatment. Therefore, for a substantial proportion of patients receiving a nasal dose in the early phases of the Pilot study (Pilot part A), the final nasal PD measurement was often within the first week post treatment. However all but 2 patients in Part A having a nasal PD at day 14 also had a PD on day 28, and therefore for the purposes of comparisons all patients with tests at pre, day 6, day 14 and day 28 were included (n=10)(section 5.2.4.7.1). Following up patients post day 28 on the basis of a suggestion of chloride secretion identified in real time during testing may be acceptable in this open label study, but would be problematic for blinding in placebo controlled trials.

5.4.4 Absence of association between nasal PD and molecular results

We did not see an association between nasal PD responses and the detection of CFTR on mRNA from post-dosing nasal brushings (section 5.2.6.1). Several possible explanations have been considered: a) this may be a genuine reflection of different levels of sensitivity of the two assays. Very few mRNA positives were detected over the course of the trial and data from other trials suggest the sensitivity of this assay is poor; b) the two sites of sampling in the nose are different, brushing being performed more proximally than nasal PD; perhaps delivery of the nasal dose was more successful further into the nose; c) since only a single post dosing nasal brushing was performed, the optimal time for detection may have been missed. The latter seems unlikely as in several cases, a negative mRNA sample was obtained on the same visit as a nasal PD with evidence of a chloride response (section 5.2.6.1).

5.4.5 Influence on chloride response

Consistent with the Lancet study, the most pronounced changes post gene therapy were observed during the ZCA phase rather than the isoprenaline phase of the nasal PD tests (35). Alton et al stated that there was a significant difference in PD response to a low chloride solution between the active and placebo groups when assessed over the duration of the measurements. No changes were detected in basal PD, response to amiloride or response to isoprenaline perfusion between active and placebo treated patients. The

absence of a placebo group in the Pilot study makes similar comparisons impossible, but we observed changes in chloride secretion in individual patients but only trends towards improved chloride secretion when results were collated (section 5.2.4.4-5.2.4.7).

5.4.6 Influence on basal PD

In the Pilot study, we did not detect any change in basal PD measurement between pre- and post treatment tests (section 5.2.4.3). No clinical trial of CF gene therapy has changed the basal PD, which is thought largely to reflect sodium transport. Importantly though, there is evidence from the literature that this may not be fundamental for clinical efficacy. Phase II studies using nasal PD as an outcome measure for the small molecule drug Ivacaftor (Vertex, US) found it had only a small, non-consistent effect on basal PD despite a significant impact on chloride secretion within subjects (97). This drug has been shown to result in clinical improvement, providing evidence for the theory that correction of the sodium transport defect in CF (responsible for the higher baseline in comparison to healthy controls) may not be necessary for clinical improvement. It will be of great interest in the future multidose trial to see if the basal PD changes following repeated dosing, and to map this on to clinical change in the cohort of patients who will have nasal PD testing during the trial.

5.4.7 Evidence of prolonged gene expression

It became apparent during the Pilot study that for some individuals, the PD changes persisted far longer than had been anticipated on the basis of pre-clinical investigation (e.g. patient 022, section 5.2.4.5). Therefore the timing of post treatment PD was amended to reflect the pre-clinical data and based on previous clinical trial (35). In Part B, PDs were performed on day 6, 12 and 28 post gene therapy. Nasal brushing was carried out at day 28. This later time point for molecular testing was chosen because of the theoretical risk of an unobtainable nasal PD post brushing because of a disrupted nasal epithelium. Absence of an intact epithelium would be expected to, in much the same way as an upper respiratory tract infection, cause an unobtainable or low basal PD. However, where a day 28 PD appeared to show a 'response', further PD measurements were undertaken (if the patient agreed) despite the fact that

brushing had already been undertaken. We learnt over the course of the trial that it was in fact possible to obtain nasal PD measurements post nasal brushing, likely reflecting the different sampling distance into the nose attained by the nasal PD catheter. That nasal brushing did not appear to contraindicate subsequent nasal PD testing is useful information to any future early phase gene therapy trial incorporating molecular and functional evidence of *CFTR* correction.

Interestingly in the Pilot study there was no single time point at which a chloride ‘response’ consistent with efficacy was most likely to be seen (section 5.2.4.4-5.2.4.7). Therefore in terms of trial planning, these data do not help to rationalise the protocol for appropriate time points for measurement. The data would lend some support for the focus to be on later time points (after 1 week). There is a balance between identifying group changes which may be small if a minority of the group ‘respond’ vs. individual changes where the majority of patients may have readings unchanged from pre-treatment values but some have impressive changes out to prolonged time points; the latter is, I believe, the situation in this trial.

Whereas for some patients maximal zero chloride secretion appeared to occur relatively early (patients 1SD011 and 017), for others it was later (patients 1SD041, 022), and for some the signal actually appeared intermittently (e.g. 1SD030), discussed below. There are additional challenges in our dataset introduced by time points of testing: patient 1SD003 only had one post dose measurement performed on day 2, which was >5mV greater in terms of ZCA delta than either pre-dosing measurement (Figure 5.5, section 5.2.4.5), and therefore no comment can be made about duration of the persistence of response.

5.4.8 Intermittent PD responses

Whereas for certain patients there was a clear evolution in nasal PD response over time between pre- and post dosing measurements, in others the response had a less obvious chronological pattern. Zero chloride responses in patient 011 and 022 provide examples where there is a clear, biologically plausible relationship between pre- and post-dosing measurements which seems unlikely

to have arisen by chance (Figure 5.5 and 5.7). In contrast, it is a challenge to explain the nasal PD results in patients with intermittent responses (e.g. patient 030). A single application of *CFTR* gene therapy to the nasal airway with plasmid DNA would not be expected to produce fluctuations in CFTR expression. If gene transfer has been successful and functional CFTR protein produced, it would be hypothesised that nasal PD should be able to detect a consistent change reflecting chloride secretion until a non-reversible end to *CFTR* gene expression. The most likely explanation for this in my view is that despite the nasal PD catheters being marked to ensure that the distance into the nose was consistent between tests, a degree of rotation may have altered the exact site of PD readings. Alternatively, reversible changes in airway integrity, for example in the presence of a subclinical viral infection, could lead to a measurement with apparently absent chloride secretion being flanked by measurements apparently demonstrating correction. If these intermittent responses are real and do reflect functional CFTR, it may mean that some patients in whom we saw an absence of response at day 28 following a response at day 14 may have had detectable changes towards normality at later time points. Therefore it is possible that more patients may have had prolonged evidence of chloride secretion indicating sustained presence of functional CFTR protein.

Intermittent nasal PD responses are not unique to the Pilot study. In the Lancet study active treatment group, electrical evidence of chloride secretion significantly greater than pre-treatment values were shown at day 1, 5 and 15 after treatment, but not at other post dosing time points (day 3,7,11 and 23) (35). Given that these time points at which a statistical difference was present were interspersed by intervals with no difference, this provides some support for the intermittent nature of the ‘signal’ seen in some Pilot study patients.

5.4.9 Electrical changes and clinical correlates

Airway PD was included as an outcome measure in the Pilot study to act as a ‘readout’ of function of *de novo* CFTR, and provide information on duration of expression. The clinical relevance of any change after our gene therapy product is unknown in several regards, for example: a) the degree of electrical correction required to influence the clinical picture and conversely b) whether the absence of change in electrical measurements correlates with a likely lack of

clinical response. Comparison of lower airway, upper airway and clinical end points from the Pilot study suggest that this relationship is likely to be complicated. However, there is some evidence that the degree of chloride secretion or basal PD measurement in the upper airway is related to clinical status. In the literature, there is evidence that the spectrum of disease caused by mutations in *CFTR* is broadly associated with the degree of abnormality detected in airway PD measurement (82).

Encouragingly, data from the UK CF Gene Therapy Consortium investigating molecular (mRNA) expression in mice suggests that gene expression increases with repeated doses of the CF gene therapy product (unpublished). Therefore absence of a change in airway PD following a single nasal dose may be inappropriate to use as a screening tool for detecting patients as ‘responders’ or ‘non-responders’ in CF gene therapy trials.

It must also be remembered that nasal PD in its current form lacks the ability to be able to quantify change in terms of proportion of cells in which electrical correction may have occurred due to *CFTR* expression. It was disappointing that the novel approach to PD testing described in section 5.3 and discussed below was unsuccessful in this regard.

5.4.10 *Novel approach to nasal PD*

Although changes in the chloride secretion response demonstrated some evidence of functional correction post gene therapy, the degree to which this correction has occurred (in terms of the proportion of cells successfully transfected) remained largely unknown using our standard experimental set up.

Using the novel set up with the LR5 unit, the author has shown that it is possible to manipulate the PD reading by applying an electrical load in the form of an additional ‘path’ for the circuit with the influence on PD dependent on the resistance load associated with this clamp (Figure 5.10 and 5.11). Tests have also demonstrated that this change in voltage is not purely due to alteration in mV caused by high impedance of the electrode / Foley catheter set up used for standard nasal PD testing (Figure 5.13).

However, unfortunately the results obtained from testing with the LR5 were not as originally hypothesised. The change in PD as a result of applying different loads was remarkably consistent over time, both when repeated in immediate sequence of testing and when performed on differing days (as observed in the CF patients' repeated measurements pre- and post CF gene therapy). Although it was consistently achievable to apply an electrical load that was able to alter the measured transepithelial PD, the change in PD both at the start and end of clamp (i.e. return to baseline PD) occurred with greater speed than anticipated.

As shown with multiple testing, in nasal measurement, return to pre-clamp PD was almost instantaneous as viewed within the proprietary LR5 software (Figure 5.10). Significant overlap was observed both within an individual and between individuals. In terms of original hypothesis that there may be a difference in rate of 'recovery' to baseline PD post clamp between CF and non-CF patients, this was not observed. This may represent a true non-difference between the two groups, or suggest that any differences are not able to be observed with the existing LR5 experimental set up. There are several reasons why no difference between CF and non-CF subjects was seen with the LR5 approach. Potentially, a) the speed at which the millivolt data is captured may not be fast enough to 'detect' any such differences; b) the maximal electrical load permissible with the LR5 set up is too 'weak'; or c) the maximal time for which the clamp could be applied (10 seconds) is not long enough to induce any measurable electrical change. In terms of the latter two points, it may be that although these parameters are sufficient to alter the transepithelial PD, they are not sufficient to produce a sustained impact on the epithelial cells in a similar manner to which is achieved by e.g. amiloride with the standard PD set up. Discussion with the inventor of the LR5 suggested that neither enhanced frequency of data capture nor altering maximal loads applied would be feasible. Therefore, although very small in terms of patient numbers and testing carried out, in the author's opinion, the lack of further investigation with the existing LR5 set up was justified. Interestingly, although further investigation of electrical manipulation may not be warranted given the disappointing results reported, an alternative approach for future investigation could include analysis of return to baseline PD following e.g. amiloride perfusion. At present, it is the amount (delta) by which the PD changes from basal level following perfusion with Ringers amiloride solution that is measured. No attempt is made to measure how

long it takes for this ‘chemical’ manipulation of PD to return to baseline values. The action of amiloride on airway epithelial PD is rapid (<2 minutes) (74). In conventional PD analysis, the outcome of interest post amiloride perfusion is maximal change in PD from baseline. Within CF subjects, rate of return to basal PD following exposure of the nasal epithelium to agents capable of altering the measured PD has been studied in the context of CF therapeutics development. For example, a study of the effects of a nasal spray to deliver benzamil (a benzyl substituted analogue of amiloride) and amiloride on nasal PD in CF involved 10 subjects and tracked nasal PD at sequential intervals until 8 hours post administration. It found that 2 hours post administration there was little difference from pre-treatment values for amiloride (in contrast to benzamil in which return to basal PD had still not occurred by 8 hours post testing) (125). Further investigation of chemical ‘loads’ could, as a result of the airway epithelial cells being arranged as a parallel electrical ‘circuit’, attempt to ‘drain’ the cells electrically during load application and then measure electrical ‘recovery’ as ion transport normalises.

5.4.11 Chapter summary and conclusions

The nasal PD results from the Pilot study provide evidence of the potential to partially restore *CFTR* function following a single nasal dose of gene therapy. However, the results were not universal, and no associations were seen with the molecular *CFTR* results. The multidose trial is designed to investigate the effect of repeated doses of gene therapy on clinical efficacy and it will be interesting to correlate these with nasal PD outcomes, and investigate how responses may be sustained or even accumulate with repeated dosing.

We must find out what the clinical consequences of certain degrees of electrical correction mean, and develop improved methods in quantifying ‘response’ in addition to being able to select the most appropriate timings for post-dosing measurements. In particular, it will be of great interest to find out whether repeated dosing (as opposed to the single nebulised dose administered in the Pilot study) results in a cumulative and more impressive change in the electrical characteristics of airway cells in CF patients receiving CF gene therapy. Ultimately, it is not measurements of the electrical components of the airway surface which are important for therapeutic development, clinical improvement

is required. Yet for subgroups of patients in these larger later phase trials, an opportunity exists to increase much needed understanding of measurement variation and functional efficacy which can then be appropriately mapped with clinical course.

Trials such as the UK CF Gene Therapy Consortium multidose trial, along with phase III trials of small molecule drugs aimed at correcting the underlying defect in *CFTR* (which also have existing data on nasal PD outcomes from earlier phase trials) will provide an unrivalled source of data to be able to map clinical change with electrical change.

6 Results: Lower airway PD measurement

6.1 Pilot study bronchoscopies

Twenty four bronchoscopies were carried out as part of the Pilot study, between March and October 2009. This included 14 pre-dose and 10 post- dose bronchoscopies. The mean age of patients that underwent bronchoscopy was 26.1 (range 17.1-44.0) years. Of the four patients with only a pre-dose bronchoscopy, two entered a post-dose ‘non bronchoscopic’ arm of the trial (this was for reasons unrelated to the bronchial PD however fortuitously, these happened to be the two patients whose pre-dose bronchoscopies were characterised by prohibitively high offsets and unstable traces). Two others withdrew from the study prior to receiving the dose of gene therapy for personal reasons.

Ten patients had paired pre- and post- gene therapy bronchoscopies (for the latter, n=4 at day 6, n=6 at day 14).

6.2 Lower airway PD

All distal lower airway PD measurements were made on a single side. With the exception of two pre- and three post- dose bronchoscopies, all distal measurements were made within airways of the right lung.

Technically acceptable readings were achieved in all patients with paired bronchoscopies. However, there was evidence of significant hyperpolarisation in response to zero chloride isoprenaline solution on the pre-dosing bronchoscopies from n=2 patients; the degree of change placed them outside 2 standard deviations of the mean for a) the rest of this cohort and b) the CF subjects included in any of our previous studies (35;113). We therefore considered it reasonable to perform data analysis both with and without these outliers. Basal measurements from both carinal and distal sites were analysed with the entire data set.

6.2.1 Agreement between JD and GD

In total, the pre- and post- dosing PDs from the ten patients with paired tests resulted in sixty sections of analysis (e.g. 20 carinal basal sections; 20 distal basal sections etc).

A flow chart summarising the outcome of the review of GD and JD values is shown in Figure 6.1. 10% of sections were discrepant following initial analysis, all were from pre-dose bronchoscopies. Following review, a consensus opinion was reached for all tests.

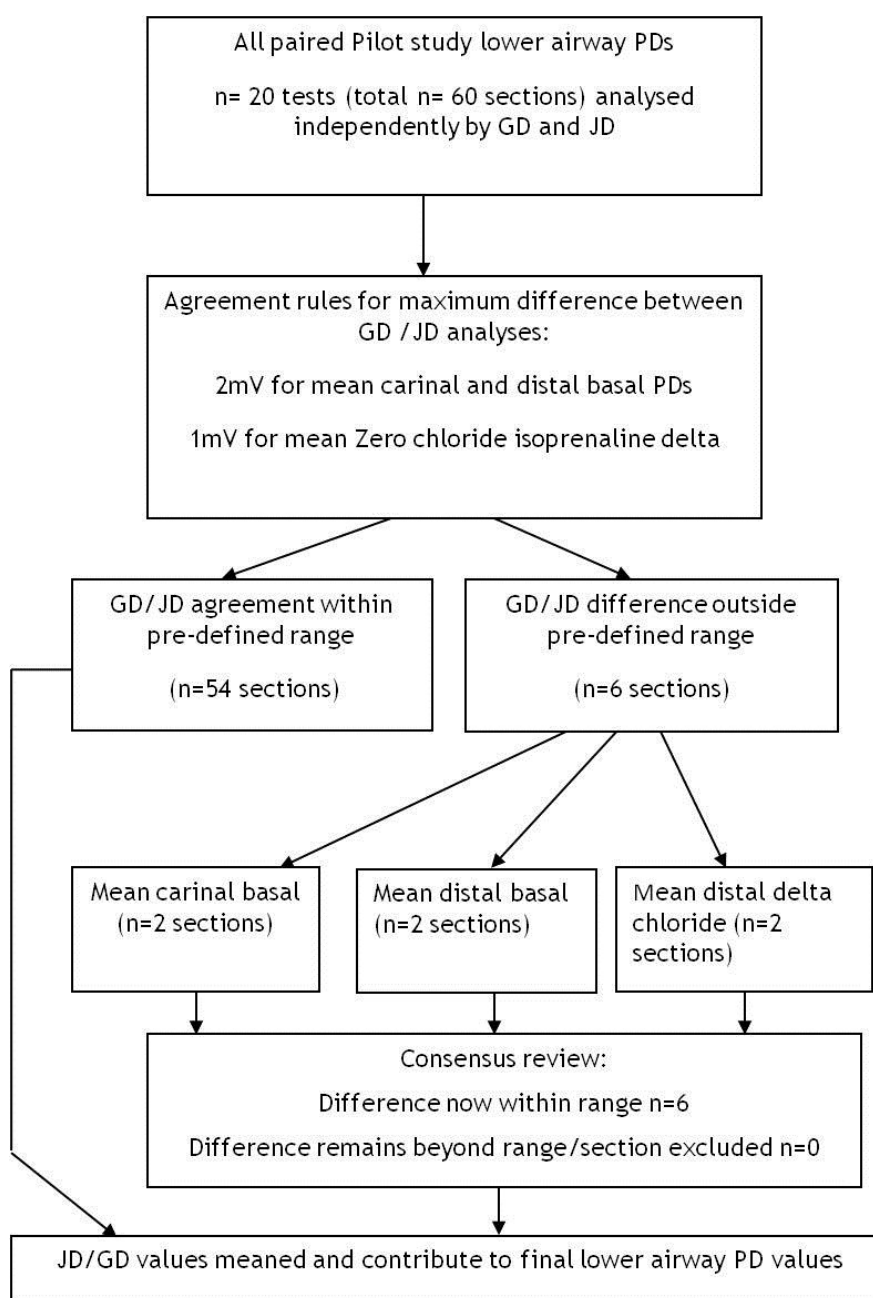


Figure 6.1. Data analysis algorithm for Pilot study lower airway PD results

To compare intra-reporter agreement for data analysis, intraclass correlation coefficients (ICC) were calculated. ICCs following initial data analysis and following consensus opinion between GD and JD are shown in Table 6.1. The agreement between analysers was superior to that observed with nasal PD analysis (Chapter 5, section 5.2.2).

	Mean carinal basal PD	Mean distal basal PD	Mean zero chloride delta
Independent analysis (initial)	0.971	0.992	0.992
Final values (following consensus review)	0.998	0.998	0.997

Table 6.1. Intra class correlation coefficients (ICC). For GD/JD bronchial PD analysis for initial independent analysis (top row) and following consensus review (bottom row).

6.2.2 Final lower airway PD results

The variability in basal carinal and distal results at the single pre- and post gene therapy time point is summarised graphically in Figure 6.2. The mean coefficients of variation for pre- and post-dose carinal basal measurements were 29.4% and 31.8% respectively. Similarly for distal basal measurements, they were 37.8% and 24.8%. Due to the presence of both negative and positive values for magnitude of chloride response, coefficient of variation was invalid. However, the variability in these measures can be appreciated visually (Figure 6.3).

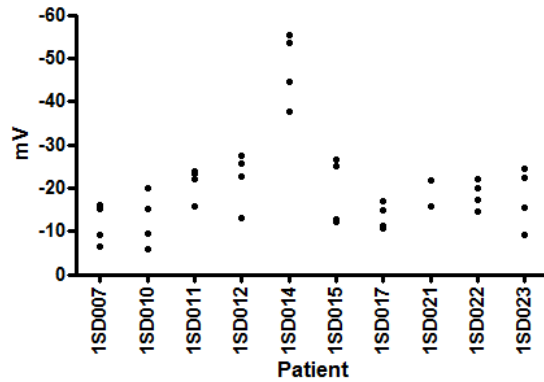


Figure 6.2A. Pre-dose basal carina

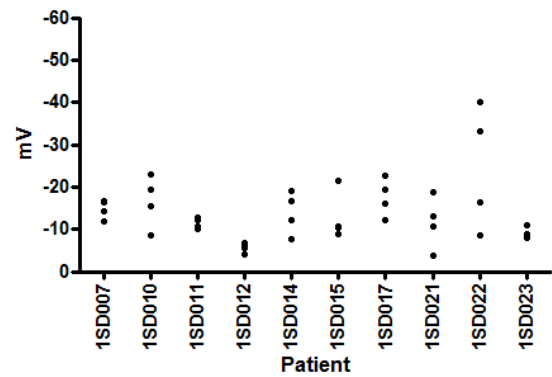


Figure 6.2B. Post-dose carina

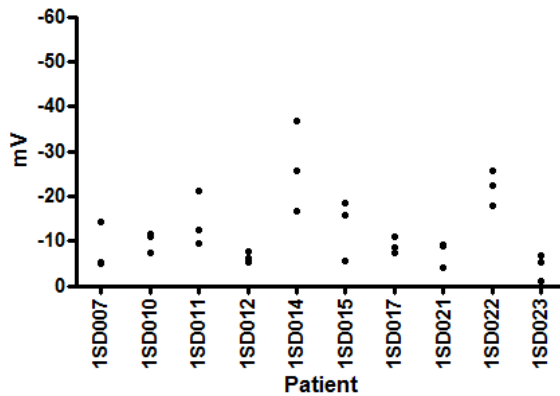


Figure 6.2C. Pre-dose distal basal

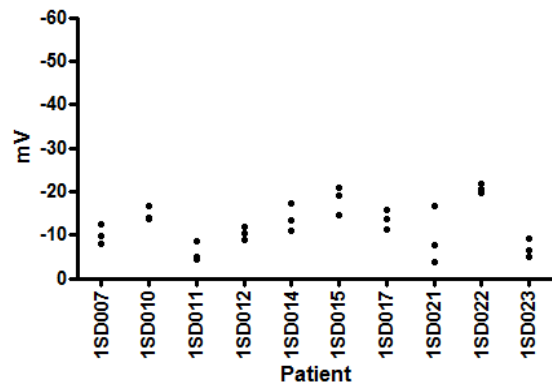


Figure 6.2D. Post-dose distal basal

Figure 6.2. Intra-patient variability in basal PD in the lower airway. The intra-patient variability in carinal basal measurements (millivolts, mV) at pre- and post- dose bronchoscopy is shown in graph 6.2A and 6.2B respectively. Similarly, variability in pre- and post- distal basal measurements is shown in graphs 6.2C and 6.2D. Each patient is represented according to study number on the X-axis, with data points for the 4 carinal (A&B) or 3 distal (C&D) measurements plotted.

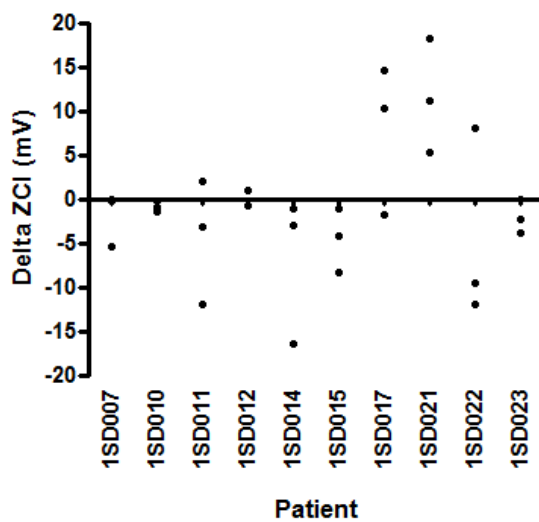


Figure 6.3A. Pre-dose chloride transport

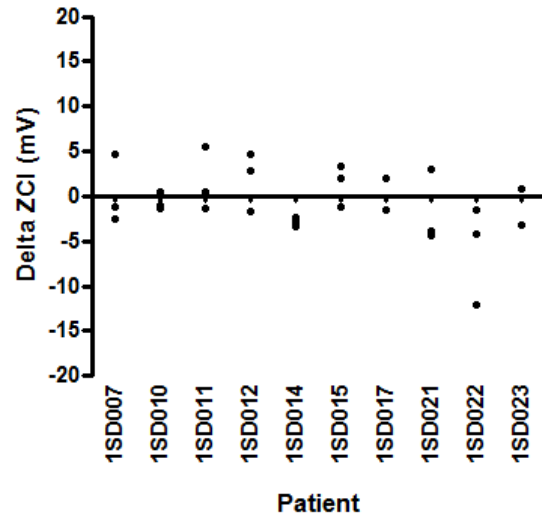


Figure 6.3B Post-dose chloride transport

Figure 6.3. Intra-patient variability in distal zero chloride isoprenaline (ZCI) response. The intra-patient variability in distal chloride transport (with deltas in millivolts, mV), at pre- (6.3A) and post- (6.3B) dose bronchoscopy is displayed. As previously, each patient is represented according to study number on the X-axis, with data points for each of the 3 distal perfusion deltas plotted.

The pre- and post dose carinal measurements at each site of measurement are summarised in Figure 6.4. The mean basal proximal measurements were on average 9.1mV higher (more negative) than those measured distally (-18.6 vs. -9.5mV, $p = 0.006$). This difference was only seen pre-dosing. A trend was seen towards lower mean carinal basal measurements post dose (Figure 6.5), making basal measurements more uniform across proximal and distal sites. No differences were observed between pre- and post-dose mean distal basal measurements (Figure 6.5).

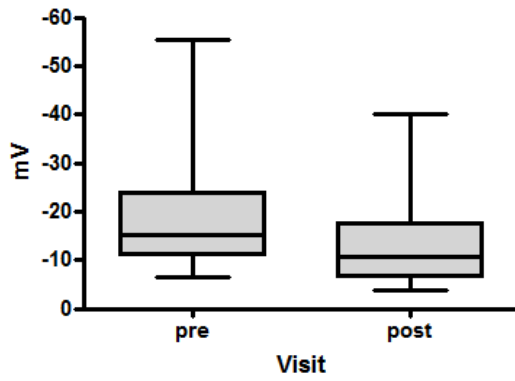


Figure 6.4A Anterior wall of carina

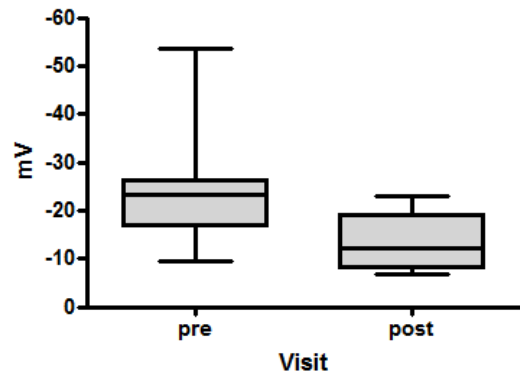


Figure 6.4B. Posterior wall of carina

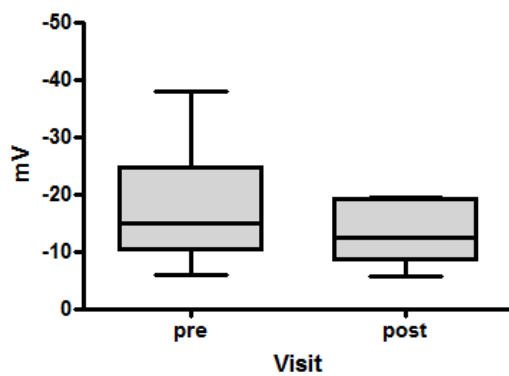


Figure 6.4C. Left lateral wall carina

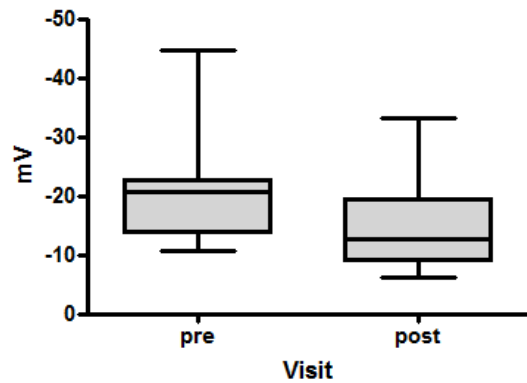


Figure 6.4D. Right lateral wall carina

Figure 6.4. Site-specific basal carinal measurements. Box and Whisker summary plots for basal carinal measurements at each site for all patients (6.4A-D) with pre- and post dose bronchoscopy (n=10), plots represent range, median and quartiles at pre- and post dose bronchoscopy visits. There were no significant differences between pre- and post dose measurements (Wilcoxon signed rank tests, $p > 0.05$ for all comparisons).

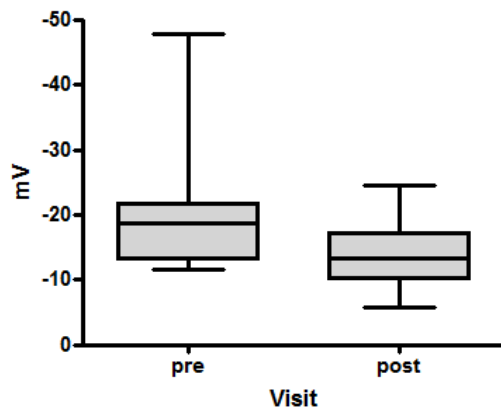


Figure 6.5A. Mean basal carinal

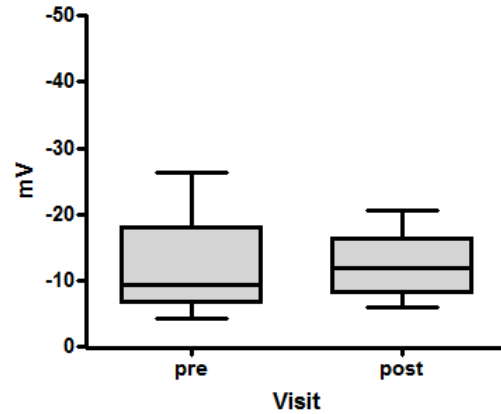


Figure 6.5B. Mean basal distal

Figure 6.5. Mean basal carinal and distal measurements. Mean basal PD measurements for all patients with pre- and post dose bronchoscopy (n=10) at carinal (6.5A) and distal (6.5B) locations in the lower airway. As above, Box and Whisker plots represent range, median and quartiles at pre- and post gene therapy bronchoscopy visits. Wilcoxon signed rank test for carinal basal measurements pre- and post dose $p=0.08$, and for basal PD at distal locations $p=0.49$.

The mean change in PD following zero chloride perfusion for the three distal perfusion sections for n=8 subjects excluding the outliers is shown in Figure 6.6. Using the same format to graphically display results as basal measurements, they are displayed as individual paired values for pre / post tests and overall group results in Box and Whisker plots. Chloride secretion was greater in 7 out of 8 patients in post- as compared to pre-dose bronchoscopies, being on average 2.44 mV greater than on pre-dosing bronchoscopies (paired t-test $p=0.03$).

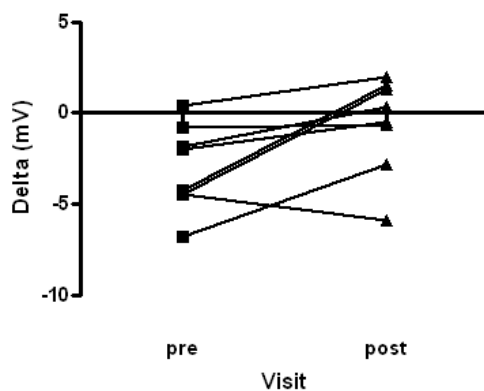


Figure 6.6A

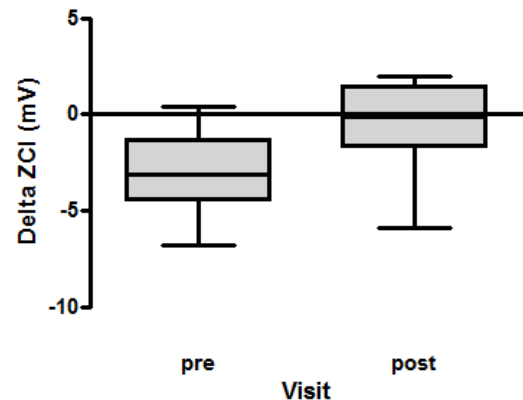


Figure 6.6B

Figure 6.6. Zero chloride responses for n=8 patients excluding the 2 outliers. Results are represented as individual patient data (Figure 6.6A) and collectively summarised (Figure 6.6B) in the Box and Whisker plots. Paired t-test for pre- vs. post measurements $p=0.03$.

The results for all 10 patients including outliers are shown in Figure 6.7. There was no statistically significant difference between pre- and post dose measures (paired t-test $p=0.93$).

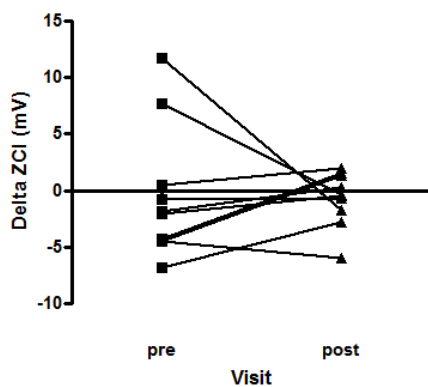


Figure 6.7A

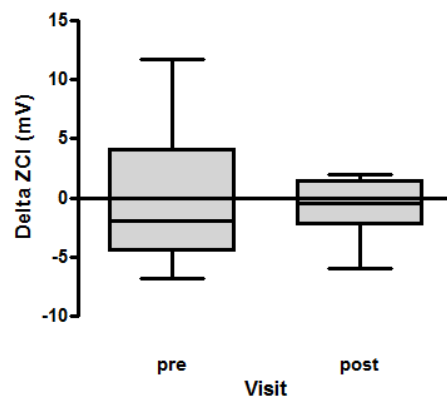


Figure 6.7B

Figure 6.7. Zero chloride responses for all n=10 patients. Results are represented as individual patient data (Figure 6.7A) and collectively summarised (Figure 6.7B) in the Box and Whisker plot. Paired t-test for pre- vs. post measurements $p=0.93$.

The trend to a reduction in basal PD at the carinal site was unexpected as no previous trial of gene therapy has succeeded in reducing sodium hyperabsorption. An alternative explanation could be that this resulted from inflammation. An intra-individual correlation of the change in this measure and that in distal chloride secretion after dosing (which is highly likely the result of gene transfer) seemed an appropriate way to explore this. For the reasons above, I limited this analysis to the 8 patients excluding the outliers; there was no significant correlation (Figure 6.8).

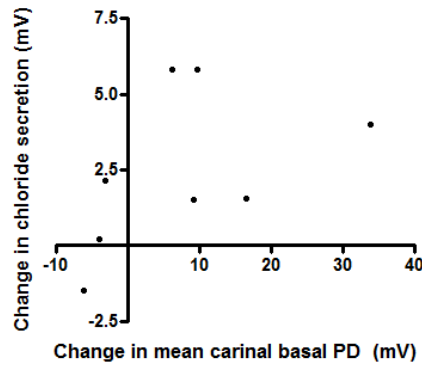


Figure 6.8. Correlation between pre- and post dose bronchoscopy measurement of chloride secretion with mean carinal basal PD. N=8 subjects, each represented by an individual data point, $r=0.49$, $p=0.22$.

There was no obvious difference between the magnitude of change in chloride secretion between pre and post-dosing tests carried out at either day 6 or day 14 post dosing. No statistical tests have been carried out as the n numbers are small, however substantial overlap was seen between results for the two time points. In the two patients with the greatest change in chloride secretion between pre and post measurements, one had their bronchoscopy carried out at day 6 (5.82 mV, patient 1SD015) and the other at day 14 (5.81 mV, patient 1SD011) post dose.

6.3 Comparison with nasal PD results

The magnitude of any changes seen in lower airway PDs was smaller than those observed for nasal measurements (Chapter 5) on an individual patient basis. There was no clear association between ‘response’ on nasal PD and change toward normal values in lower airway PD. The two patients with the highest mV achieved in response to zero chloride amiloride perfusion on nasal PD testing were 1SD011 and 1SD022 (see Chapter 5, section 5.2.4.5 and 5.2.4.7.2). Interestingly, 1SD011 had the second highest increase in chloride secretion in the lower airway post dose (day 14). In contrast, 1SD022 had no evidence of increased chloride secretion in the lower airway post dose (day 14). 1SD022 was the only subject in the cohort of 8 patients to have less chloride secretion on post dosing measurements. However, this patient also had a significantly lower basal nasal PD at day 14 than at any of their other nine nasal PD measurements

(basal PD = -11.6mV at day 14 visit, median basal PD from all 10 tests throughout Pilot study = -29.0mV), perhaps indicating a contributory effect from inflammation.

6.4 Comparison with molecular *CFTR* results

As described in Chapter 2, section 2.7, following the pre- and post-dose bronchoscopy lower airway PD measurements, bronchial brushings were taken to determine molecular evidence of effective gene transfer including vector-specific *CFTR* mRNA. All of the pre-dosing samples were (as expected) negative, as were the majority of the post-dosing samples; only two patients (1SD007 and 1SD023) had vector-specific *CFTR* mRNA detected in post dosing bronchial brushings (one at day 6 and the other at day 14) as described in Chapter 4, section 4.8.1. Furthermore, although above limits of detection, the proportion of *CFTR* mRNA was not quantifiable for one of these patients. Both these patients had improved chloride secretion post dose (2.15 and 1.51 mV respectively). However the n=5 other patients with lower airway PD results consistent with improved chloride secretion had negative lower airway mRNA results.

6.5 Discussion

In Chapter 2, section 2.6, a method for measuring lower airway PD using pharmaceutical grade reagents and incorporating steps to evaluate both basal and chloride secretory responses to perfusion has been reported, and successfully used to collect all lower airway PD measurements throughout the Pilot study. The results obtained from pre- and post dosing measurements have been reported in this chapter, alongside associations between these and molecular evidence of *de novo* CFTR expression and to upper airway PD results. Several themes merit further discussion.

6.5.1 *Experimental set up and design*

6.5.1.1 Development

The methodology described in Chapter 2, section 2.6 has allowed data collection pre- and post a single dose of CF gene therapy in the Pilot study and is currently being utilised in the ongoing multidose trial. A newly designed circuit and careful circuit preparation, in conjunction with warmed solutions and other modifications enabled technically acceptable recordings.

However, lower airway PD has its limitations. These include technical difficulty, time for testing and requirement for general anaesthetic. The test itself adds time (approximately 30 minutes) to the bronchoscopy procedure whilst the patient is under general anaesthetic - we observed one SAE following bronchoscopy as described in Chapter 4, section 4.9. Time for data analysis at conclusion of trial was also not insignificant. Similar to nasal PD data analysis, each test took the author around 30 minutes to analyse, not including time taken to collate results and reach a consensus opinion with JD on any discrepant results as defined above.

6.5.1.2 Comparison with literature

Development of a method to measure lower airway PD suitable for use in a clinical trial in the current regulatory environment was a primary aim of this

work. A single lumen catheter technique for measuring lower airway PD, similar to that developed for the Pilot study, has previously been described by our group for testing suitable for use with the paediatric bronchoscope (113). This was no longer suitable as GCP requirements have evolved. Davies et al measured basal PD in CF and non-CF subjects, with tracheal values being significantly more negative in CF than non-CF (113). Lower basal PDs were observed at more distal sites, where there was no longer any difference between CF and non-CF children. In the paediatric study, responses to ZCI perfusion in a segmental (3rd generation) bronchus in the non-CF group showed a mean hyperpolarisation of 10.1mV after 3 minutes perfusion, but in contrast a small depolarisation was observed in the CF group (mean change -1.3mV) with no overlap between groups. Comparison data for adults measured using a double lumen catheter technique were reported in the same paper, with mean CF adult tracheal and segmental bronchus PD measurements around -17mV (vs. -9mV for control subjects), and mean sub segmental and wedge PDs in CF adults around -15mV and -6mV respectively. All measurements by Davies et al (113) were obtained using the glass Calomel electrodes and intravenous electrical bridge, rather than the solid Ag/AgCl electrodes utilised in the Pilot study.

A solid Ag/AgCl electrode technique has been described by Fajac et al (114), with a thin wire extending down the bronchoscope with a small ball at its distal end providing electrical contact. Using their electrode set up (which permits basal readings but no perfusion protocol), Fajac et al, first validated the technique against nasal PD readings with a perfused electrode and Ag/AgCl electrode in the nose, before using the solid electrode for measurements in the lower airways. This was used to propose a normal range for lower airway basal PDs, although of note their subject group involved adult patients with lung cancer and interstitial lung disease (mean (SD) age 60 (19) years) rather than healthy controls. Despite potential limitations caused by patient group, their results are consistent with other studies reported in the literature, with a small but progressive decline in basal PD traversing distally through the lower airway from trachea to segmental and sub segmental bronchi (mean PD distal trachea -13.4mV; mean PD segmental and sub segmental bronchi -10.6mV) (113;114).

6.5.1.3 Suggestions for improvement

Although we were satisfied with the technical quality of the lower airway PD measurements in the Pilot study, several suggestions can be made for further refinement of the experimental set up. The method developed for Pilot used the sheath of a cytology brush to contain the fluid acting as an electrical conductor for measurements through the bronchoscope (Chapter 2, section 2.6.2.5). This tube is ideal in that it is sterile, fits easily through the lumen of the bronchoscope channel and is designed and marketed for use during bronchoscopy. The process of simply discarding the cytology brush was wasteful, but we were unable to source an alternative type of tubing with the relevant clinical and regulatory approval.

6.5.2 Data analysis

Data analysis of all lower airway PDs was performed blinded to patient and test order, according to the criteria described in Chapter 2, section 2.6.1 and a strict quality control process was employed to ensure that results from the two analysers (author and JD) were appropriate. The degree of intra-reporter agreement was excellent pre- consenting with a limited number of tests with discrepant results. The author considers it vital that traces can be reviewed and discussed to reach a consensus agreement as small discrepancies in a mV result could have implications for final group comparisons and differences between pre- and post dosing tests when any potential treatment effects may be small. Due to the scarcity of research groups undertaking lower airway PD measurements on a worldwide scale, there is no accepted ‘gold standard’ approach to data analysis. However, components of scoring employed here including blinding, more than one analyser, and taking the mean of multiple data points over a short period of stability to identify the appropriate mV result should be considered standard.

The importance of multiple basal (carinal and distal) and zero chloride isoprenaline perfusion steps was demonstrated by the high coefficient of variation between measurements performed at the same bronchoscopy. This variability in PD according to site (e.g. between different 5th generation airways) means it is crucial to perform multiple measurements to obtain a more

representative ‘mean’ PD at that site. Each perfusion step takes five minutes and therefore there is an importance balance between extent of data collection and potential for additional patient morbidity due to general anaesthesia and additional fluid instilled into the airway.

6.5.3 Results from Pilot study

6.5.3.1 Basal measurements

In the CF gene therapy study described by Alton et al (the ‘Lancet study’) (35), basal lower airway PD measurements at the four carinal sites were not different from each other and therefore the mean value was used to define basal carinal PD. We utilised this approach for Pilot study results analysis. The basal measurements at the carina in the Pilot study were comparable to those reported in the Lancet study (35). Caution must be exercised in comparing absolute values due to differences in the methodology used for measurement, nevertheless the absolute values for measured PD were similar. No changes in basal PD post dose were observed in the Lancet study, however the date of testing was different (all post dosing measurements made on day 2).

In the Pilot study, despite the absence of a statistically significant difference, the post dose carinal basal PD measurements appeared to be lower (less negative) than pre-dosing measurements (Figure 6.5, section 6.2.2). As mentioned, this could be related to inflammation or evidence of functional efficacy. Unfortunately the data do not provide definitive proof of mechanism. There are factors suggesting that these measurements could be related to inflammation (e.g. a known impact on distal airways post dosing as detected by gas transfer, see Chapter 4 section 4.5.2). Although by day 6 post dose (the earliest time point for post dose bronchoscopy in the Pilot study), blood and clinical evidence of a post dose inflammatory reaction had subsided (Chapter 4 sections 4.4-4.5), it is possible that there were on-going inflammatory responses at the level of the airway epithelium at the time of post dose bronchoscopy at either day 6 or day 14. Inflammation causes the airway epithelium to become ‘leaky’ electrically and, as occurs during acute upper respiratory tract infection, PD measurement is not possible. However a greater ability of the epithelium to

secrete chloride on post -dosing tests (discussed below) may instead point more towards efficacy.

6.5.3.2 Chloride secretion

Two patients had significant evidence of chloride secretion in the lower airway at the pre-dosing bronchoscopy visit (in ≥ 2 distal carinal zero chloride isoprenaline perfusion measurements), as shown in Figure 6.3, section 6.2.2. For both patients, hyperpolarisation of PD in response to ZCI exceeded the expected range in CF (35;113). The presence of chloride secretion on pre- dosing tests (which interestingly was absent on their post dosing measurements) makes interpretation of any cohort change in overall chloride secretion challenging. Later in the Pilot study, patients were excluded from the nasal cohort if they had evidence of chloride secretion on pre-dosing nasal PDs and received a lung dose (in the non-invasive cohort, and therefore no lower airway PD measurements) alone. This approach is also being employed in the multidose trial. For completeness, the author chose to present findings both including and excluding these two subjects (1SD017 and 1SD021).

Interestingly, neither patient had significant evidence of chloride secretion on pre-dose nasal PD. Again, the relationship between upper and lower airway results is uncertain. The Pilot study was primarily a safety study and outcome measures reflecting this were obtained at bronchoscopy (including any histological evidence of change in endobronchial biopsies post dose), however if bronchoscopy is performed for the sole purpose of efficacy measurements, then patients with pre-dose evidence of chloride secretion should not undergo post-dose bronchoscopy.

For those patients without significant evidence of chloride secretion on pre-dosing tests, the improvement in chloride secretion post dosing provides some evidence to support successful *CFTR* gene transfer (Figure 6.6, section 2.2.2). The clinical consequences of this small but statistically significant improvement are unknown but will be investigated in the multidose trial which is taking place currently. Any therapeutic consequences of potentially being able to alter proximal basal measurements are also unknown, particularly as pathologically, CF is a disease of the distal airways. Statistically significant changes in mV in the lower airway suggestive of improved chloride secretion were reported by Alton

et al in the Lancet study for change from pre-dose measurements at day 2 post dose in their gene therapy clinical trial (35). Again caution must be used to compare absolute values, the magnitude of change in mV seen in the Pilot study was smaller than that observed in the Lancet study, where an improvement of around 5mV in the active treatment group was reported (which the authors describe as representing 25% restoration of chloride channel function in comparison to healthy controls) (35). Due to considerable differences in measurement circuit and PD solutions it is difficult to say with certainty whether the results are truly different or whether for the same patient the same results would be obtained using different testing set-up. In addition, as stated previously the time points for post-dosing measurements were not the same.

In the Pilot study, no post dosing bronchoscopies were carried out after day 14. However, changes in nasal PD measurements were sometimes seen at later time points. In addition, patients with a small increment in chloride secretion may have had a greater degree of chloride secretion at day 14 (or later) but the single post -dosing measurement time point makes any comment on change within individuals over time (i.e. potential duration of gene expression and *de novo* functional CFTR) not possible. The clinical trial plasmid used in the Pilot study was known from *in vitro* work to have a slower onset but longer duration of action than the plasmid investigated previously (100). Therefore it is possible that our opportunity to detect efficacy was impinged by the cross-sectional time points chosen. However subjecting a patient to repeated general anaesthetics for serial lower airway PDs would be unethical and likely contribute to additional morbidity. This remains a challenge for planning clinical trials incorporating lower airway PD measurements.

6.5.3.3 Comparison with upper airway PD and molecular evidence of CFTR

As described in section 6.3, there was no relationship between evidence of chloride secretion on post dose bronchoscopy and similar evidence on nasal PD. The reasons behind this are likely to include differences in method of delivery of the nasal and lower airway dose, the site of deposition and the sites of measurement. There was also no association between lower airway PD and mRNA positive results for vector specific CFTR in post dosing tests (section 6.4). Although seven patients (70%) had greater chloride secretion on post dosing

tests, the magnitude of this increase as measured in millivolts was small. This, in combination with the low number of patients positive on mRNA testing makes any formal attempt at sensitivity analysis futile, however the results from the Pilot study suggest that lower airway PD is more sensitive than mRNA. There were no PD 'negative' and mRNA 'positive' results.

The nasal PD findings from the Pilot study (Chapter 4) could have provided a useful benchmark for the most appropriate time points for lower airway PD, however the absence of a clear association between upper and lower airway results, and maximal chloride secretion at different time points post dosing between different patients again creates additional challenge. Furthermore the absence of a relationship between molecular evidence of *CFTR* function (mRNA) and lower airway PD results is also not helpful to guide appropriate time points for optimal changes of detecting maximal change in future clinical trial.

With the uncertainties in interpretation of results and, in particular, the relationship with other outcome measures explored in the Pilot study, does the author feel that lower airway PDs are an ethically justified outcome measure in early phase clinical trials? On balance, the author suggests that the information obtained from lower airway measurement can, in combination with other outcome measures, provide additional information at the site of interest. The difficulty is that it may not be possible to determine the significance (or otherwise) of the measurements until they are mapped to change in clinically meaningful end points as explored in a multidose trial. This would be useful because once we understand how certain PD changes correlate clinically, we can utilise this information to rationalise clinical trial design.

6.5.4 Chapter summary and conclusions

The lower airway measurements performed in the Pilot study lend support to evidence of CFTR expression in the lower airway, although the magnitude of electrical change was small and less than previously reported in the Lancet study (35). We will need to await the results of the current multidose trial in order to be able to make real sense of the results described here. The same measurement set up described here is being taken forward for use in the multidose trial. Until further information becomes available, the lower airway

PD results should not be reviewed in isolation but taken in combination with upper airway results, and molecular evidence of successful *CFTR* gene transfer and functional CFTR protein.

7 Overall summary and conclusions

This thesis has described the investigation of outcome measures for CF gene therapy clinical trials. It has concentrated on two categories: those that focus on the ion transport underlying defect in *CFTR* and physiological tests which measure downstream consequences of disease. These outcome measures have been evaluated within the context of the UKCFGTC clinical programme.

The opportunity for timely evaluation of results from the observational Run-In study alongside the interventional Pilot study provided an excellent platform for data analysis. Each results chapter in this thesis has concluded with a discussion section relevant to the issues raised therein. This final chapter brings together the themes which traverse the thesis.

7.1 Measuring outcome

7.1.1 *Why?*

Accurate measurement and appropriate interpretation of outcome is fundamental for the evaluation of novel treatments. This thesis has described well-established, evolving, and novel outcome measures. However, even for longstanding outcome measures such as FEV₁ issues relating to appropriate handling of raw data have been shown critically to affect interpretation of the tests performed (from meeting inclusion criteria for a trial to influencing assessment of rate of decline over time). Others such as lung clearance index are more recent and still in a relative infancy in terms of recognition in a clinical trial and regulatory setting. The relationship between commercial companies and academia in terms of opportunities for analysis of the true raw data and appropriate interpretation has again been a learning point. Informed scientific interpretation relies fundamentally on acquiring accurate raw data and being able to convert these to meaningful values to permit appropriate data analysis. The advent of increasingly sophisticated measurement devices is a tempting setting for extensive data collection. An ability to translate this into something relevant for a clinical trial (or the wider clinical picture) remains a further, equally important and certainly not insignificant, challenge. The author now has

a greater awareness of the potential problems resulting from ‘black box’ tests which will be relevant to future research and clinical practice.

7.1.2 When?

The most appropriate time points for data collection will differ according to both the intervention and the outcome of interest. This was illustrated in the Pilot study with the safety outcomes (including airway physiology and systemic inflammatory markers) having differential post-dose time points for maximal change. This has implications for protocol design including efforts to minimise the number of study visits/ assays performed at each visit without compromising the evaluation of appropriate outcomes. It is not always possible to be aware of optimal time points at the stage of protocol design, particularly for first in man studies such as the Pilot study. However, recognisable patterns in post-intervention responses can enable a more pragmatic approach to subsequent data collection in future trials as is being applied in our multidose trial. Even more challenging was the finding that there was no uniform time point of maximal change for evidence of functional efficacy in the nasal PD tests from the Pilot study - cross-sectional testing at intermittent study intervals may still ‘miss’ a response in some cases. The importance of this will be clearer following the multidose trial, when downstream endpoints can more appropriately be correlated with functional evidence of *CFTR* on nasal (and/or) lower airway PD.

The Run In study and Pilot study have demonstrated the potential utility of increasing the number of tests undertaken prior to any intervention rather than relying on a single pre-dose measurement for comparisons. For tests with high intra-subject variability during periods of clinical stability, repeated measurements improve the reliability of a pre-dose result on which to base post dose comparisons. However, the influence of inherent variability for post dosing tests should also not be forgotten, particularly when comparisons with a single post dose value are made.

7.1.3 How?

The Run-In and Pilot study data support the hypothesis that the lower airway requires more than one measure of function or status to reflect pathologies at

different sites. Outcomes for efficacy or toxicity may also require the researcher to 'look' at a different site of the airway to those conventionally measured in CF. The analysis of airway physiology outcomes in the Run-In study demonstrated that conventionally held impressions of particular indices reflecting a certain airway site may not hold true. The finding that airway physiological outcomes correlated with measures of systemic inflammation in the Run-In study (but interestingly not in the Pilot study post dosing) may reflect a single or parallel underlying disease process yet at present we do not have appropriate tests to help disentangle this further. The complex relationship between disease severity and variability between results at different testing dates (as seen with the Run-In study dataset) is also a challenge for effective clinical trial design and evaluation.

The Pilot study was primarily a safety study but also included *CFTR* molecular and functional end points. The author supports this combined approach due to the relatively small patient population and logistical, time and financial implications of multiple separate clinical trials. Although no relationships were observed in Pilot between molecular or functional *CFTR* outcomes with downstream end points, the multidose trial is powered to detect clinical benefit and therefore will provide an opportunity to analyse in detail the consequences of 'positive' or 'response' results at endpoints detecting correction of the underlying molecular defect. This will also permit greater understanding of the results obtained during the single dose study. There are likely to be issues with sensitivity of PD and molecular testing (mRNA); the absence of change in airway PD does not necessarily translate into the absence of clinical benefit (conversely change in airway PD does not necessarily correlate with clinical improvements). It remains a challenge how best to use functional and molecular evidence of *CFTR* correction in terms of interpretation and application to later phase gene therapy clinical trials. However there is evidence in the wider literature involving small molecule drugs that changes in nasal PD can relate to clinical endpoints (99). Such comparisons will also help determine the most appropriate method of defining and quantifying any response seen.

The author has described modification of the conventional set up for lower airway PD measurement and successfully used this to measure bronchial PD pre and post a single nebulised dose of CF gene therapy in the Pilot study. Serial

nasal PD measurements have also been made, with subsequent detailed and intensive data analysis. Due to a limitation of existing airway PD measurements to quantify the proportion of cells corrected by gene therapy, a novel set up was proposed with the theoretical capability to achieve this. It was disappointing that the new approach to PD measurement, using the LR5, did not move the field forward in terms of being able to quantify change.

7.2 The multidose trial

The Run-In study was designed to ‘run in’ to the multidose trial. However, it became apparent from data analysis that during periods of clinical stability, there is often no discernible slope representing decline in respiratory status across relatively short longitudinal follow up. Despite this, the primary and secondary outcomes of interest were identified following analysis of results in the Run-In study. Optimal dosing strategy was guided by the Pilot study results.

The placebo controlled multidose trial of nebulised pGM169/GL67A began after the author left the department and is ongoing at the time of writing. It has been referred to on multiple occasions throughout this thesis, in part reflecting the close association between all three clinical studies (Run In, Pilot and multidose) as described in Chapter 1, section 1.5. Patient recruitment commenced in June 2012, following completion of the Pilot study and a successful extensive pre-clinical toxicology package. Patients from London and Edinburgh sites are randomised to active treatment or placebo administered on a monthly basis for one year. The CF gene therapy complex used in the multidose trial is identical to that used in the Pilot study (pGM169/GL67A), with a 5ml nebulised dose. The placebo is 5ml 0.9% sodium chloride. All patients receive planned adjuvant treatment with paracetamol for a 24 hour period on dosing days. A sub-cohort of patients are also randomised to additional repeated (monthly) nasal dosing with associated nasal PD measurement, and to an invasive cohort where bronchoscopies are performed to assess molecular and functional evidence of successful *CFTR* gene transfer after repeated dose (or placebo) administration using the methodology developed by the author described in Chapter 2, section 2.6.

The multidose trial will allow an appropriately powered evaluation of clinical efficacy. The author also hopes it will enable us to investigate further the relationship between functional measurements and evidence of clinical efficacy. It will also be particularly interesting to discover whether the apparent improvement in LCI between pre- and post dosing measurements in the Pilot study is confirmed with greater patient numbers and whether any such effect is sustained or improved further with repeated dosing.

7.3 The future

There is no ideal single outcome measure for CF because it is a heterogeneous disease with a non-uniform path of disease progression. Processes occurring as an adverse response post therapeutic intervention may be separate from usual disease processes (and affect different parts of the airway) and therefore require assessment of outcomes not typically associated with monitoring of CF lung disease.

We are entering a new era in the management of cystic fibrosis, with a genuine prospect of managing our patients with treatments aimed at correcting the underlying defect rather than fire fighting downstream consequences of disease. It is of fundamental importance that we continue to investigate better ways of measuring disease, both for efficacy and toxicity, and ensure that such measures are able to be mapped on to meaningful outcomes for patients. With new technologies, this field will continue to evolve. Just as commitment to novel therapies and their translation from the laboratory to the clinic is a research priority, so must there be ongoing commitment to measuring outcome. A ‘fine-tooth comb’ approach to development and appropriate evaluation of emerging candidates, whilst not forgetting simultaneously to put existing outcome measures through forensic examination, is mandatory for success.

8 References

- (1) VanDevanter D, Konstan MW. Outcome measures for clinical trials assessing treatment of cystic fibrosis lung disease. *Clin Invest* 2012;2:163-75.
- (2) Rosenfeld M. An overview of endpoints for cystic fibrosis clinical trials: one size does not fit all. *Proc Am Thorac Soc* 2007;4(4):299-301.
- (3) Davies JC, Alton EW, Bush A. Cystic fibrosis. *BMJ* 2007;335(7632):1255-9.
- (4) Correlation between genotype and phenotype in patients with cystic fibrosis. The Cystic Fibrosis Genotype-Phenotype Consortium. *N Engl J Med* 1993;329(18):1308-13.
- (5) Picard E, Aviram M, Yahav Y, Rivlin J, Blau H, Bentur L, et al. Familial concordance of phenotype and microbial variation among siblings with CF. *Pediatr Pulmonol* 2004;38(4):292-7.
- (6) Gabriel SE, Clarke LL, Boucher RC, Stutts MJ. CFTR and outward rectifying chloride channels are distinct proteins with a regulatory relationship. *Nature* 1993;363:263-8.
- (7) Schreiber R, Greger R, Nitschke R, Kunzelmann K. Cystic fibrosis transmembrane conductance regulator activates water conductance in *Xenopus* oocytes. *Pflugers Arch* 1997;434:841-7.
- (8) Mall M, Kunzelmann K, Hipper A, Busch AE. cAMP stimulation of CFTR-expressing *Xenopus* oocytes activates a chromanol-inhibitable K conductance. *Pflugers Arch* 1996;432:516-22.
- (9) Engelhardt JF, Yankaskas JR, Ernst SA, Yang Y, Marino CR, Boucher RC, et al. Submucosal glands are the predominant site of CFTR expression in the human bronchus. *Nat Genet* 1992;2(3):240-8.
- (10) Engelhardt JF, Zepeda M, Cohn JA, Yankaskas JR, Wilson JM. Expression of the cystic fibrosis gene in adult human lung. *J Clin Invest* 1994;93(2):737-49.
- (11) Knowles MR, Boucher RC. Mucus clearance as a primary innate defense mechanism for mammalian airways. *J Clin Invest* 2002;109(5):571-7.
- (12) Matsui H, Grubb BR, Tarran R, Randell SH, Gatzky JT, Davis CW, et al. Evidence for periciliary liquid layer depletion, not abnormal ion composition, in the pathogenesis of cystic fibrosis airways disease. *Cell* 1998;95(7):1005-15.
- (13) Dodge JA, Lewis PA, Stanton M, Wilsher J. Cystic fibrosis mortality and survival in the UK: 1947-2003. *Eur Respir J* 2007;29(3):522-6.

- (14) Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, et al. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science* 1989;245(4922):1066-73.
- (15) Koshland DE, Jr. The Cystic Fibrosis Gene Story. *Science* 1989;245(4922):1029.
- (16) Griesenbach U, Alton EW. Gene transfer to the lung: lessons learned from more than 2 decades of CF gene therapy. *Adv Drug Deliv Rev* 2009;61(2):128-39.
- (17) Alton EW, Middleton PG, Caplen NJ, Smith SN, Steel DM, Munkonge FM, et al. Non-invasive liposome-mediated gene delivery can correct the ion transport defect in cystic fibrosis mutant mice. *Nat Genet* 1993;5(2):135-42.
- (18) Drumm ML, Pope HA, Cliff WH, Rommens JM, Marvin SA, Tsui LC, et al. Correction of the cystic fibrosis defect in vitro by retrovirus-mediated gene transfer. *Cell* 1990;62(6):1227-33.
- (19) Zabner J, Couture LA, Gregory RJ, Graham SM, Smith AE, Welsh MJ. Adenovirus-mediated gene transfer transiently corrects the chloride transport defect in nasal epithelia of patients with cystic fibrosis. *Cell* 1993;75(2):207-16.
- (20) Caplen NJ, Alton EW, Middleton PG, Dorin JR, Stevenson BJ, Gao X, et al. Liposome-mediated CFTR gene transfer to the nasal epithelium of patients with cystic fibrosis. *Nat Med* 1995;1(1):39-46.
- (21) Crystal RG, McElvaney NG, Rosenfeld MA, Chu CS, Mastrangeli A, Hay JG, et al. Administration of an adenovirus containing the human CFTR cDNA to the respiratory tract of individuals with cystic fibrosis. *Nat Genet* 1994;8(1):42-51.
- (22) Halbert CL, Standaert TA, Aitken ML, Alexander IE, Russell DW, Miller AD. Transduction by adeno-associated virus vectors in the rabbit airway: efficiency, persistence, and readministration. *J Virol* 1997;71(8):5932-41.
- (23) Johnson LG, Olsen JC, Sarkadi B, Moore KL, Swanstrom R, Boucher RC. Efficiency of gene transfer for restoration of normal airway epithelial function in cystic fibrosis. *Nat Genet* 1992;2(1):21-5.
- (24) Johnson LG, Boyles SE, Wilson J, Boucher RC. Normalization of raised sodium absorption and raised calcium-mediated chloride secretion by adenovirus-mediated expression of cystic fibrosis transmembrane conductance regulator in primary human cystic fibrosis airway epithelial cells. *J Clin Invest* 1995;95(3):1377-82.
- (25) Accurso FJ, Rowe SM, Clancy JP, Boyle MP, Dunitz JM, Durie PR, et al. Effect of VX-770 in persons with cystic fibrosis and the G551D-CFTR mutation. *N Engl J Med* 2010;363(21):1991-2003.

- (26) Rosenecker J, Harms KH, Bertele RM, Pohl-Koppe A, Mutius E, Adam D, et al. Adenovirus infection in cystic fibrosis patients: implications for the use of adenoviral vectors for gene transfer. *Infection* 1996;24(1):5-8.
- (27) Walters RW, Grunst T, Bergelson JM, Finberg RW, Welsh MJ, Zabner J. Basolateral localization of fiber receptors limits adenovirus infection from the apical surface of airway epithelia. *J Biol Chem* 1999;274(15):10219-26.
- (28) Song Y, Lou HH, Boyer JL, Limberis MP, Vandenberghe LH, Hackett NR, et al. Functional cystic fibrosis transmembrane conductance regulator expression in cystic fibrosis airway epithelial cells by AAV6.2-mediated segmental trans-splicing. *Hum Gene Ther* 2009;20(3):267-81.
- (29) Sinn PL, Arias AC, McCray PB. Repeated administration of lentiviral vector to mouse nasal epithelia enhances expression without emergence of blocking mucosal immune responses. *Pediatr Pulmonol* 2007;42(S30):285.
- (30) Mitomo K, Griesenbach U, Inoue M, Somerton L, Meng C, Akiba E, et al. Toward Gene Therapy for Cystic Fibrosis Using a Lentivirus Pseudotyped With Sendai Virus Envelopes. *Mol Ther* 2010;18(6):1173-1182.
- (31) Hacein-Bey-Abina S, Von KC, Schmidt M, McCormack MP, Wulffraat N, Leboulch P, et al. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 2003;302(5644):415-9.
- (32) Hacein-Bey-Abina S, Garrigue A, Wang GP, Soulier J, Lim A, Morillon E, et al. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J Clin Invest* 2008;118(9):3132-42.
- (33) Pringle IA, Hyde SC, Gill DR. Non-viral vectors in cystic fibrosis gene therapy: recent developments and future prospects. *Expert Opin Biol Ther* 2009;9(8):991-1003.
- (34) Simoes S, Filipe A, Faneca H, Mano M, Penacho N, Duzgunes N, et al. Cationic liposomes for gene delivery. *Expert Opin Drug Deliv* 2005;2(2):237-54.
- (35) Alton EW, Stern M, Farley R, Jaffe A, Chadwick SL, Phillips J, et al. Cationic lipid-mediated CFTR gene transfer to the lungs and nose of patients with cystic fibrosis: a double-blind placebo-controlled trial. *Lancet* 1999;353(9157):947-54.
- (36) Hyde SC, Southern KW, Gileadi U, Fitzjohn EM, Mofford KA, Waddell BE, et al. Repeat administration of DNA/liposomes to the nasal epithelium of patients with cystic fibrosis. *Gene Ther* 2000;7(13):1156-65.
- (37) Schwartz DA, Quinn TJ, Thorne PS, Sayeed S, Yi AK, Krieg AM. CpG motifs in bacterial DNA cause inflammation in the lower respiratory tract. *J Clin Invest* 1997;100(1):68-73.
- (38) Hemmi H, Takeuchi O, Kawai T, Kaisho T, Sato S, Sanjo H, et al. A Toll-like receptor recognizes bacterial DNA. *Nature* 2000;408(6813):740-5.

- (39) McLachlan G, Stevenson BJ, Davidson DJ, Porteous DJ. Bacterial DNA is implicated in the inflammatory response to delivery of DNA/DOTAP to mouse lungs. *Gene Therapy* 2000;7(5):384-92.
- (40) Yew NS, Zhao H, Wu IH, Song A, Tousignant JD, Przybylska M, et al. Reduced inflammatory response to plasmid DNA vectors by elimination and inhibition of immunostimulatory CpG motifs. *Mol Ther* 2000;1(3):255-62.
- (41) Hyde SC, Pringle IA, Abdullah S, Lawton AE, Davies LA, Varathalingam A, et al. CpG-free plasmids confer reduced inflammation and sustained pulmonary gene expression. *Nat Biotechnol* 2008;26(5):549-51.
- (42) Davies JC, Alton EW. Gene therapy for cystic fibrosis. *Proc Am Thorac Soc* 2010;7(6):408-14.
- (43) Moss RB, Rodman D, Spencer LT, Aitken ML, Zeitlin PL, Waltz D, et al. Repeated adeno-associated virus serotype 2 aerosol-mediated cystic fibrosis transmembrane regulator gene transfer to the lungs of patients with cystic fibrosis: a multicenter, double-blind, placebo-controlled trial. *Chest* 2004;125(2):509-21.
- (44) Moss RB, Milla C, Colombo J, Accurso F, Zeitlin PL, Clancy JP, et al. Repeated aerosolized AAV-CFTR for treatment of cystic fibrosis: a randomized placebo-controlled phase 2B trial. *Hum Gene Ther* 2007;18(8):726-32.
- (45) Crook K, McLachlan G, Stevenson BJ, Porteous DJ. Plasmid DNA molecules complexed with cationic liposomes are protected from degradation by nucleases and shearing by aerosolisation. *Gene Ther* 1996;3(9):834-9.
- (46) Davis PB, Byard PJ, Konstan MW. Identifying treatments that halt progression of pulmonary disease in cystic fibrosis. *Pediatr Res* 1997;41(2):161-5.
- (47) Gustafsson PM, Aurora P, Lindblad A. Evaluation of ventilation maldistribution as an early indicator of lung disease in children with cystic fibrosis. *Eur Respir J* 2003;22(6):972-9.
- (48) Horsley AR, Gustafsson PM, Macleod KA, Saunders C, Greening AP, Porteous DJ, et al. Lung clearance index is a sensitive, repeatable and practical measure of airways disease in adults with cystic fibrosis. *Thorax* 2008;63(2):135-40.
- (49) Aurora P, Bush A, Gustafsson P, Oliver C, Wallis C, Price J, et al. Multiple-breath washout as a marker of lung disease in preschool children with cystic fibrosis. *Am J Respir Crit Care Med* 2005;171(3):249-56.
- (50) Davies JC, Sheridan HS, Lee PSL, Song TS. Lung clearance index to evaluate the effect of Ivacaftor on lung function in subjects with CF who have the G551D CFTR mutation and mild lung disease. *Thorax* 2012;67(Suppl 2):A58.

- (51) Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J* 2005;26(2):319-38.
- (52) McFadden ER, Jr., Linden DA. A reduction in maximum mid-expiratory flow rate. A spirographic manifestation of small airway disease. *Am J Med* 1972;52(6):725-37.
- (53) Hamutcu R, Rowland JM, Horn MV, Kaminsky C, MacLaughlin EF, Starnes VA, et al. Clinical findings and lung pathology in children with cystic fibrosis. *Am J Respir Crit Care Med* 2002;165(8):1172-5.
- (54) Goris ML, Zhu HJ, Blankenberg F, Chan F, Robinson TE. An automated approach to quantitative air trapping measurements in mild cystic fibrosis. *Chest* 2003;123(5):1655-63.
- (55) Bakker EM, Volpi S, Salonini E, van der Wiel-Kooij EC, Sintnicolaas CJ, Hop WC, et al. Improved treatment response to dornase alfa in cystic fibrosis patients using controlled inhalation. *Eur Respir J* 2011;38(6):1328-35.
- (56) Burgel PR. The role of small airways in obstructive airway diseases. *Eur Respir Rev* 2011;20(119):23-33.
- (57) Sorkness RL, Bleecker ER, Busse WW, Calhoun WJ, Castro M, Chung KF, et al. Lung function in adults with stable but severe asthma: air trapping and incomplete reversal of obstruction with bronchodilation. *J Appl Physiol* 2008;104(2):394-403.
- (58) Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. *Am J Respir Crit Care Med* 1999;159(1):179-87.
- (59) Knudson RJ, Slatin RC, Lebowitz MD, Burrows B. The maximal expiratory flow-volume curve. Normal standards, variability, and effects of age. *Am Rev Respir Dis* 1976;113(5):587-600.
- (60) Morris JF, Koski A, Johnson LC. Spirometric standards for healthy nonsmoking adults. *Am Rev Respir Dis* 1971;103(1):57-67.
- (61) Quanjer PH, Tammeling GJ, Cotes JE, Pedersen OF, Peslin R, Yernault JC. Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J Suppl* 1993;16:5-40.
- (62) Rosenthal M, Bain SH, Cramer D, Helms P, Denison D, Bush A, et al. Lung function in white children aged 4 to 19 years: I--Spirometry. *Thorax* 1993;48(8):794-802.
- (63) Quanjer PH, Hall GL, Stanojevic S, Cole TJ, Stocks J. Age- and height-based prediction bias in spirometry reference equations. *Eur Respir J* 2012;40(1):190-7.
- (64) Stanojevic S, Wade A, Stocks J, Hankinson J, Coates AL, Pan H, et al. Reference ranges for spirometry across all ages: a new approach. *Am J Respir Crit Care Med* 2008;177(3):253-60.

- (65) Cole TJ. The LMS method for constructing normalized growth standards. *Eur J Clin Nutr* 1990;44(1):45-60.
- (66) Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. *Stat Med* 1992;11(10):1305-19.
- (67) Gustafsson PM, Kallman S, Ljungberg H, Lindblad A. Method for assessment of volume of trapped gas in infants during multiple-breath inert gas washout. *Pediatr Pulmonol* 2003;35(1):42-9.
- (68) Horsley AR, Gustafsson PM, Macleod KA, Saunders C, Greening AP, Porteous DJ, et al. Lung clearance index is a sensitive, repeatable and practical measure of airways disease in adults with cystic fibrosis. *Thorax* 2008;63(2):135-40.
- (69) Amin R, Subbarao P, Jabar A, Balkovec S, Jensen R, Kerrigan S, et al. Hypertonic saline improves the LCI in paediatric patients with CF with normal lung function. *Thorax* 2010;65:379-83.
- (70) Amin R, Subbarao P, Lou W, Jabar A. The effect of dornase alfa on ventilation inhomogeneity in patients with cystic fibrosis. *Eur Respir J Suppl* 2011;37:806-12.
- (71) Anderson MP, Gregory RJ, Thompson S, Souza DW, Paul S, Mulligan RC, et al. Demonstration that CFTR is a chloride channel by alteration of its anion selectivity. *Science* 1991;253(5016):202-5.
- (72) Stutts MJ, Canessa CM, Olsen JC, Hamrick M, Cohn JA, Rossier BC, et al. CFTR as a cAMP-dependent regulator of sodium channels. *Science* 1995 ;269(5225):847-50.
- (73) Boucher RC, Stutts MJ, Knowles MR, Cantley L, Gatzky JT. Na⁺ transport in cystic fibrosis respiratory epithelia. Abnormal basal rate and response to adenylate cyclase activation. *J Clin Invest* 1986;78(5):1245-52.
- (74) Knowles M, Gatzky J, Boucher R. Increased bioelectric potential difference across respiratory epithelia in cystic fibrosis. *N Engl J Med* 1981;305(25):1489-95.
- (75) Middleton PG, Geddes DM, Alton EW. Protocols for in vivo measurement of the ion transport defects in cystic fibrosis nasal epithelium. *Eur Respir J* 1994;7(11):2050-6.
- (76) Chen JH, Stoltz DA, Karp PH, Ernst SE, Pezzulo AA, Moninger TO, et al. Loss of anion transport without increased sodium absorption characterizes newborn porcine cystic fibrosis airway epithelia. *Cell* 2010 ;143(6):911-23.
- (77) Itani OA, Chen JH, Karp PH, Ernst S, Keshavjee S, Parekh K, et al. Human cystic fibrosis airway epithelia have reduced Cl⁻ conductance but not increased Na⁺ conductance. *Proc Natl Acad Sci U S A* 2011;108(25):10260-5.
- (78) Knowles MR, Carson JL, Collier AM, Gatzky JT, Boucher RC. Measurements of nasal transepithelial electric potential differences in normal human subjects in vivo. *Am Rev Respir Dis* 1981;124(4):484-90.

- (79) Alton EW, Hay JG, Munro C, Geddes DM. Measurement of nasal potential difference in adult cystic fibrosis, Young's syndrome, and bronchiectasis. *Thorax* 1987;42(10):815-7.
- (80) Alton EW, Currie D, Logan-Sinclair R, Warner JO, Hodson ME, Geddes DM. Nasal potential difference: a clinical diagnostic test for cystic fibrosis. *Eur Respir J* 1990;3(8):922-6.
- (81) Wilschanski M, Dupuis A, Ellis L, Jarvi K, Zielenski J, Tullis E, et al. Mutations in the cystic fibrosis transmembrane regulator gene and in vivo transepithelial potentials. *Am J Respir Crit Care Med* 2006 ;174(7):787-94.
- (82) Leal T, Fajac I, Wallace HL, Lebecque P, Lebacq J, Hubert D, et al. Airway ion transport impacts on disease presentation and severity in cystic fibrosis. *Clin Biochem* 2008 Jul;41(10-11):764-72.
- (83) De Boeck K, Wilschanski M, Castellani C, Taylor C, Cuppens H, Dodge J, et al. Cystic fibrosis: terminology and diagnostic algorithms. *Thorax* 2006;61(7):627-35.
- (84) Ahrens RC, Standaert TA, Launspach J, Han SH, Teresi ME, Aitken ML, et al. Use of nasal potential difference and sweat chloride as outcome measures in multicenter clinical trials in subjects with cystic fibrosis. *Pediatr Pulmonol* 2002;33(2):142-50.
- (85) McCarty NA, Standaert TA, Teresi M, Tuthill C, Launspach J, Kelley TJ, et al. A phase I randomized, multicenter trial of CPX in adult subjects with mild cystic fibrosis. *Pediatr Pulmonol* 2002;33(2):90-8.
- (86) Flotte TR, Zeitlin PL, Reynolds TC, Heald AE, Pedersen P, Beck S, et al. Phase I trial of intranasal and endobronchial administration of a recombinant adeno-associated virus serotype 2 (rAAV2)-CFTR vector in adult cystic fibrosis patients: a two-part clinical study. *Hum Gene Ther* 2003;14(11):1079-88.
- (87) Gill DR, Southern KW, Mofford KA, Seddon T, Huang L, Sorgi F, et al. A placebo-controlled study of liposome-mediated gene transfer to the nasal epithelium of patients with cystic fibrosis. *Gene Ther* 1997 ;4(3):199-209.
- (88) Knowles MR, Hohneker KW, Zhou Z, Olsen JC, Noah TL, Hu PC, et al. A controlled study of adenoviral-vector-mediated gene transfer in the nasal epithelium of patients with cystic fibrosis. *N Engl J Med* 1995;333(13):823-31.
- (89) Porteous DJ, Dorin JR, McLachlan G, vidson-Smith H, Davidson H, Stevenson BJ, et al. Evidence for safety and efficacy of DOTAP cationic liposome mediated CFTR gene transfer to the nasal epithelium of patients with cystic fibrosis. *Gene Ther* 1997;4(3):210-8.
- (90) Zabner J, Couture LA, Gregory RJ, Graham SM, Smith AE, Welsh MJ. Adenovirus-mediated gene transfer transiently corrects the chloride transport defect in nasal epithelia of patients with cystic fibrosis. *Cell* 1993;75(2):207-16.

- (91) Zabner J, Ramsey BW, Meeker DP, Aitken ML, Balfour RP, Gibson RL, et al. Repeat administration of an adenovirus vector encoding cystic fibrosis transmembrane conductance regulator to the nasal epithelium of patients with cystic fibrosis. *J Clin Invest* 1996;97(6):1504-11.
- (92) Zabner J, Cheng SH, Meeker D, Launspach J, Balfour R, Perricone MA, et al. Comparison of DNA-lipid complexes and DNA alone for gene transfer to cystic fibrosis airway epithelia in vivo. *J Clin Invest* 1997;100(6):1529-37.
- (93) Konstan MW, Davis PB, Wagener JS, Hilliard KA, Stern RC, Milgram LJ, et al. Compacted DNA nanoparticles administered to the nasal mucosa of cystic fibrosis subjects are safe and demonstrate partial to complete cystic fibrosis transmembrane regulator reconstitution. *Hum Gene Ther* 2004;15(12):1255-69.
- (94) De Boeck K, Kent L, Davies J, Derichs N, Amaral M, Rowe S, et al. CFTR biomarkers: time for promotion to surrogate endpoint? *Eur Respir J* 2012;41(1);203-16.
- (95) Goldman MJ, Yang Y, Wilson JM. Gene therapy in a xenograft model of cystic fibrosis lung corrects chloride transport more effectively than the sodium defect. *Nat Genet* 1995 Feb;9(2):126-31.
- (96) Kerem E, Hirawat S, Armoni S, Yaakov Y, Shoseyov D, Cohen M, et al. Effectiveness of PTC124 treatment of cystic fibrosis caused by nonsense mutations: a prospective phase II trial. *Lancet* 2008;372(9640):719-27.
- (97) Accurso FJ, Rowe SM, Clancy JP, Boyle MP, Dunitz JM, Durie PR, et al. Effect of VX-770 in persons with cystic fibrosis and the G551D-CFTR mutation. *N Engl J Med* 2010;363(21):1991-2003.
- (98) Wilschanski M, Miller LL, Shoseyov D, Blau H, Rivlin J, Aviram M, et al. Chronic ataluren (PTC124) treatment of nonsense mutation cystic fibrosis. *Eur Respir J* 2011;38(1):59-69.
- (99) Ramsey BW, Davies J, McElvaney NG, Tullis E, Bell SC, Drevinek P, et al. A CFTR potentiator in patients with cystic fibrosis and the G551D mutation. *N Engl J Med* 2011;365(18):1663-72.
- (100) Hyde SC, Pringle IA, Abdullah S, Lawton AE, Davies LA, Varathalingam A, et al. CpG-free plasmids confer reduced inflammation and sustained pulmonary gene expression. *Nature Biotechnology* 2008;26(5):549-51.
- (101) McLachlan G, Davidson H, Holder E, Davies LA, Pringle IA, Sumner-Jones SG, et al. Pre-clinical evaluation of three non-viral gene transfer agents for cystic fibrosis after aerosol delivery to the ovine lung. *Gene Therapy* 2011;18(10):996-1005.
- (102) Gill DR, Nunez-Alonso GA, McLachlan G, Cheng SH, Scheule RK, Hyde SC, et al. Aerosol characteristics of DNA/lipid formulations for Gene Therapy clinical studies (abstract only). *Pediatr Pulmonol* 2008; 43:307.

- (103) Davies LA, Nunez-Alonso GA, Hebel HL, Scheule RK, Cheng SH, Hyde SC, et al. A novel mixing device for the reproducible generation of nonviral gene therapy formulations. *Biotechniques* 2010;49(3):666-8.
- (104) Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1(8476):307-10.
- (105) MacIntyre N, Crapo RO, Viegi G, Johnson DC, van der Grinten CP, Brusasco V, et al. Standardisation of the single-breath determination of carbon monoxide uptake in the lung. *Eur Respir J* 2005;26(4):720-35.
- (106) Quittner AL, Butt A, Messer MA, Modi AC, Watrous M. Development and validation of the cystic fibrosis questionnaire in the United States - A health-related quality-of-life measure for cystic fibrosis. *Chest* 2005 ;128(4):2347-54.
- (107) Modi AC, Quittner AL. Validation of a disease-specific measure of health-related quality of life for children with cystic fibrosis. *Journal of Pediatric Psychology* 2003;28(8):535-45.
- (108) Quittner AL, Modi AC, Wainwright C, Otto K, Kirihaara J, Montgomery AB. Determination of the Minimal Clinically Important Difference Scores for the Cystic Fibrosis Questionnaire-Revised Respiratory Symptom Scale in Two Populations of Patients With Cystic Fibrosis and Chronic *Pseudomonas aeruginosa* Airway Infection. *Chest* 2009;135(6):1610-8.
- (109) Rowe SM, Accurso F, Clancy JP. Detection of cystic fibrosis transmembrane conductance regulator activity in early-phase clinical trials. *Proc Am Thorac Soc* 2007;4(4):387-98.
- (110) Standaert TA, Boitano L, Emerson J, Milgram LJ, Konstan MW, Hunter J, et al. Standardized procedure for measurement of nasal potential difference: an outcome measure in multicenter cystic fibrosis clinical trials. *Pediatr Pulmonol* 2004;37(5):385-92.
- (111) Boucher RC, Jr., Bromberg PA, Gatzky JT. Airway transepithelial electric potential in vivo: species and regional differences. *J Appl Physiol* 1980 ;48(1):169-76.
- (112) Wilschanski M, Famini H, Strauss-Liviatan N, Rivlin J, Blau H, Bibi H, et al. Nasal potential difference measurements in patients with atypical cystic fibrosis. *Eur Respir J* 2001;17(6):1208-15.
- (113) Davies JC, Davies M, McShane D, Smith S, Chadwick S, Jaffe A, et al. Potential difference measurements in the lower airway of children with and without cystic fibrosis. *Am J Respir Crit Care Med* 2005;171(9):1015-9.
- (114) Fajac I, Lacronique J, Lockhart A, l'Ava-Santucci J, Dusser DJ. Silver/silver chloride electrodes for measurement of potential difference in human bronchi. *Thorax* 1998;53(10):879-81.

- (115) Boucher RC, Cotton CU, Gatzky JT, Knowles MR, Yankaskas JR. Evidence for reduced Cl⁻ and increased Na⁺ permeability in cystic fibrosis human primary cell cultures. *J Physiol* 1988;405:77-103.
- (116) Stanojevic S, Wade A, Stocks J. Reference values for lung function: past, present and future. *European Respiratory Journal* 2010;36(1):12-9.
- (117) Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, et al. Interpretative strategies for lung function tests. *Eur Respir J* 2005;26(5):948-68.
- (118) Konstan MW, Morgan WJ, Butler SM, Pasta DJ, Craib ML, Silva SJ, et al. Risk factors for rate of decline in forced expiratory volume in one second in children and adolescents with cystic fibrosis. *Journal of Pediatrics* 2007;151(2):134-9.
- (119) Taylor-Robinson D, Whitehead M, Diderichsen F, Olesen HV, Pressler T, Smyth RL, et al. Understanding the natural progression in %FEV1 decline in patients with cystic fibrosis: a longitudinal study. *Thorax* 2012;67(10):860-6.
- (120) Gustafsson PM, De Jong PA, Tiddens HAWM, Lindblad A. Multiple-breath inert gas washout and spirometry versus structural lung disease in cystic fibrosis. *Thorax* 2008;63(2):129-34.
- (121) Bankier AA, Van Muylem A, Knoop C, Estenne M, Gevenois PA. Bronchiolitis obliterans syndrome in heart-lung transplant recipients: Diagnosis with expiratory CT. *Radiology* 2001;218(2):533-9.
- (122) Lauder SN, Taylor PR, Clark SR, Evans RL, Hindley JP, Smart K, et al. Paracetamol reduces influenza-induced immunopathology in a mouse model of infection without compromising virus clearance or the generation of protective immunity. *Thorax* 2011;66(5):368-74.
- (123) Clancy JP, Rowe SM, Bebok Z, Aitken ML, Gibson R, Zeitlin P, et al. No detectable improvements in cystic fibrosis transmembrane conductance regulator by nasal aminoglycosides in patients with cystic fibrosis with stop mutations. *Am J Respir Cell Mol Biol* 2007;37(1):57-66.
- (124) Leonard A, Lebecque P, Dingemanse J, Leal T. A randomized placebo-controlled trial of miglustat in cystic fibrosis based on nasal potential difference. *J Cyst Fibros* 2012;11(3):231-6.
- (125) Rodgers HC, Knox AJ. The effect of topical benzamil and amiloride on nasal potential difference in cystic fibrosis. *Eur Respir J* 1999;14(3):693-6.

9 Appendix

9.1 Run-In study: schedule of visits and investigations

Test		Daily	Weekly	Every visit	One occasion during study visit 1-4
Symptom diary card		✓			
Clinical examination				✓	
Pulse oximetry				✓	
Sputum (spontaneous/induction if required)	Clinical microbiology			✓	
	Total and diff cell count			✓	
	24 hour collection weight			✓ (24hr pre)	
	Wet and dry weight			✓	
	Inflammatory markers			✓	
Bloods	Full blood count and differential			✓	
	C-Reactive protein			✓	
	Urea, creatinine and electrolytes,			✓	
	Liver function tests			✓	
	Inflammatory proteins			✓	
	CFTR mutation analysis (if not known)				✓
	T-Lymphocytes (anti CFTR immune responses)				✓
Spirometry	EasyOne			✓	
	Piko-6		✓	✓	
Exhaled breath condensate	pH			✓	
	Inflammatory markers			✓	
	Glucose			✓	
Urine	Cotinine			✓	
	Pregnancy test (if female and CT visit)				✓
	Inflammatory markers			✓	
Quality of life questionnaire				✓	
Lung clearance index				✓	
Shuttle exercise test				✓	
Pedometer				✓ (7 days pre-visit)	
CT chest					✓
Lung mucocilliary clearance (MCC) scan					✓

9.2 Pilot study visits and investigations

Outcome measure		Safety measure?	Efficacy measure?	Initial visit	Pre-dose bronchoscopy visit (Invasive cohort part A)	Baseline parameters (<7 days pre-dosing)	Dosing visit (*=pre, \$=post)	Day 2	Day 6	Day 14	Post dose bronchoscopy (invasive cohort part A), day 6 or 14 only	Day 28
Physical exam/ History		✓		*	*	*	* \$	*	*	*	*	*
Diary cards		✓		Daily from enrolment to end of trial (day 28)								
Routine bloods	CRP	✓		*	*		*	*	*	*	*	*
	FBC & differential	✓		*	*		*	*	*	*	*	*
	ESR	✓					*	*	*	*	*	*
	Liver function tests	✓		*	*		*	*	*	*	*	*
	Renal	✓		*	*		*	*	*	*	*	*
Specialised bloods	Coagulation	✓		*								
	cytokines	✓		*			*	*	*	*	*	*
	T-cells/ ANA	✓					*					*
Urine	routine urinalysis	✓		*			*	*			*	*
	Cotinine	✓		*								
	Exploratory			*			*	*			*	*
	Pregnancy test	✓		*			*					
Sputum (induced if required [■])	Count/ differential	✓	✓	*	only if spontaneous	*	only if spontaneous	*	*	*	only if spontaneous	*
	Culture	✓		*	only if spontaneous	*	only if spontaneous	*	*	*	only if spontaneous	*
	Cytokines	✓	✓	*	only if spontaneous	*	only if spontaneous	*	*	*	only if spontaneous	*
Radiology [■]	Limited CT	✓				*		*			*	*
Physiology	EasyOne Spirometry	✓		*	*	*	* \$	*	*	*	*	*
	Piko spirometry	✓		Daily from enrolment to end of trial (day 28)								
	TLCO	✓				*		*		*	* (if at day 14)	*
	Pulse oximetry	✓			*	*	* \$	*	*		*	*
	LCI	✓		*		*		*	(*)	(*)	Not performed	*
nasal PD			✓	*	*	*		*	(*)	(*)	*	(*)
nasal brushing	mRNA/ IHC		✓		*						*	
bronchial PD			✓		*						*	
bronchial Brushing (x10)	mRNA		✓		*						*	
bronchial biopsy (x2)	Histology	✓			*						*	

[■]Chest CT scan and sputum induction were only performed in Pilot part A.

NB. For patients receiving a Lung dose only (i.e. no nasal dose of gene therapy) in Pilot part B there was no day 6 visit. The testing dates for nasal PD depended on the patient's dose cohort. Nasal brushings for the first 3 patients receiving a 1ml nasal dose/10ml lung dose and 2ml nasal dose/20ml lung dose were performed on day 2.

9.3 Run-In study symptom score

Run-In study visits were undertaken during periods of clinical stability. This was assessed by using a symptom score. The symptom score sheet was developed by the UKCFGTC prior to the start of the Run-In study, and had been used previously in clinical studies performed by the Consortium. This allowed an assessment of symptoms as reported by the patient to determine whether they were clinically 'stable' and could participate in their Run-In study visit. The score included seven questions relating to different aspects of respiratory function:

1. How severe is your cough?
2. How severe is your breathlessness?
3. How tired or lethargic are you?
4. How far can you walk easily?
5. How much sputum are you producing?
6. Has the shade of sputum changed?
7. How thick is your sputum?

Potential responses were on a 5 point Likert scale from -2 (much worse/dark/thicker than usual) to +2 (much better than usual). No overall change from usual is represented by a score of 0. The score from each question was summed to produce a final symptom score (potential range -14 to +14). Patients were required to have a score of greater than -4 to continue with their Run-In study visit.

9.4 Composition of nasal PD solutions

Stock PD solutions are prepared at a pharmacy manufacturing unit and supplied sterile in EVA bags (100ml volume). Solutions are supplied in sealed EVA bags have a 3 month expiry when kept refrigerated. Solution composition is according to the US Cystic Fibrosis Foundation Therapeutics Development Network Standard Operating Procedure.

9.4.1 Ringers (R) solution

Raw Material	Concentration
NaCl	7.83 g/L
CaCl ₂ .2H ₂ O	0.33g/L
MgCl ₂ .6H ₂ O	0.24g/L
K ₂ HPO ₄	0.42g/L
KH ₂ PO ₄	0.05g/L

9.4.2 Ringers amiloride (RA) solution

Raw Material	Concentration
NaCl	7.83 g/L
CaCl ₂ .2H ₂ O	0.33g/L
MgCl ₂ .6H ₂ O	0.24g/L
K ₂ HPO ₄	0.42g/L
KH ₂ PO ₄	0.05g/L
Amiloride	0.03g/L

9.4.3 Zero chloride amiloride (ZCA) solution

Raw Material	Concentration
Sodium Gluconate	29.43g/L
Calcium Gluconate	0.95g/L
K ₂ HPO ₄ USP	0.42g/L
KH ₂ PO ₄	0.05g/L
MgSO ₄ .7H ₂ O	0.3g/L
Amiloride	0.03g/L

9.4.4 Zero chloride isoprenaline and Zero chloride amiloride isoprenaline solution

Isoprenaline sulphate is added to either ZC (bronchial PD) or ZCA (nasal PD) solution immediately prior to PD measurement (in view of an expiry time of 2 hours). 0.125mg is added to each 50ml ZC or ZCA volume.

9.5 Composition of bronchial PD solutions

Stock PD solutions are prepared at a pharmacy manufacturing unit and supplied sterile in EVA bags (100ml volume). Solutions are supplied in sealed EVA bags have a 3 month expiry when kept refrigerated. Solution composition is according to the US Cystic Fibrosis Foundation Therapeutics Development Network Standard Operating Procedure.

9.5.1 Ringers (R) solution

Raw Material	Concentration
NaCl	7.83 g/L
CaCl ₂ .2H ₂ O	0.33g/L
MgCl ₂ .6H ₂ O	0.24g/L
K ₂ HPO ₄	0.42g/L
KH ₂ PO ₄	0.05g/L

9.5.2 Zero chloride (ZC) solution

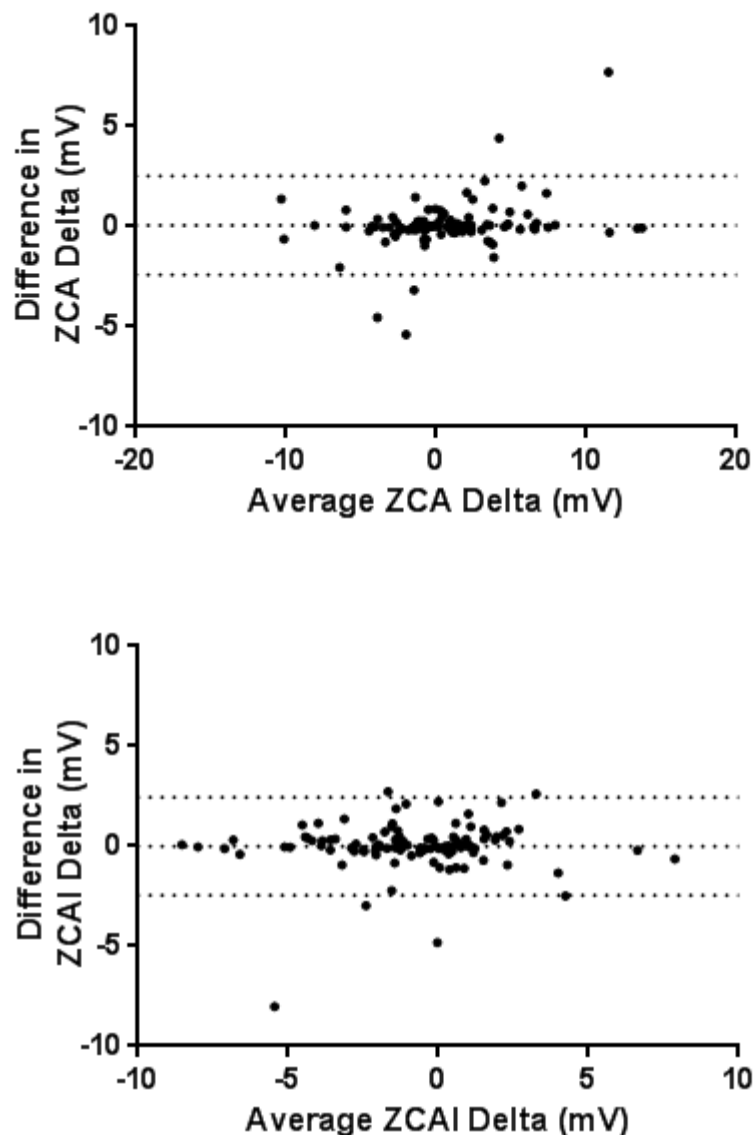
Raw Material	Concentration
Sodium Gluconate	29.43g/L
Calcium Gluconate	0.95g/L
K ₂ HPO ₄ USP	0.42g/L
KH ₂ PO ₄	0.05g/L
MgSO ₄ .7H ₂ O	0.3g/L

9.5.3 Zero chloride isoprenaline solution

Isoprenaline sulphate is added to ZC (bronchial PD) solution immediately prior to PD measurement (in view of an expiry time of 2 hours). 0.125mg is added to each 50ml ZC volume.

9.6 Nasal PD analysis agreement for chloride secretion

Nasal PD tests in the Pilot study were analysed by the author (GD) and JD. The figures below show the agreement between analysers following their own individual independent analysis, and prior to consenting to agree final results for all nasal PDs in the 2ml nasal dose cohort.



The top figure is a summary of results for chloride secretion in response to perfusion with ZCA solution (mean bias 0.06mV, 95% limits of agreement -2.41 to 2.79mV) for the analysers JD and GD. The bottom figure is a summary of chloride secretion in response to ZCAI solution (mean bias -0.01mV, 95% limits of agreement -2.45 to 2.44mV). There was no bias between analysers when assessing the magnitude of chloride secretion in Pilot nasal PD tests.