Author's response to reviews

Title: Polymorphisms in the glutathione pathway modulate cystic fibrosis severity: a cross-sectional study

Authors:

Fernando AL Marson (fernandolimarson@hotmail.com)
Carmen S Bertuzzo (bertuzzo@unicamp.br)
Antônio F Ribeiro (anferi@uol.com.br)
José D Ribeiro (jdirceuribeiro@gmail.com)

Version: 6
Date: 12 November 2013

Author's response to reviews: see over
**Manuscript:** Polymorphisms in the glutathione pathway modulate cystic fibrosis severity: a cross-sectional study

**Authors:** Fernando Augusto de Lima Marson, Carmen Sílvia Bertuzzo, Antônio Fernando Ribeiro and José Dirceu Ribeiro

The authors acknowledge the collaboration of the BMC group and the reviewers, for their contribution with criticism, suggestions and corrections made on our paper entitled "Polymorphisms in the glutathione pathway modulate cystic fibrosis severity: a cross-sectional study".

We are submitting a new version of the corrected article to the appreciation of the editor and reviewers.

**Reviewer:** Frauke Stanke

**Reviewer's report:** “Polymorphisms in the glutathione pathway modulate cystic fibrosis severity: a cross-sectional study” by de Lima Marson and colleagues.

For clarification: The reviewer understands that this is a resubmission of the data previously submitted to The Journal in a manuscript entitled “Genetic interaction of GSH metabolic pathway genes in Cystic Fibrosis?.” – Answer: Both studies analyze the same polymorphisms, however, have different questions and answers, including the model for data analysis and specific objectives. In the context of previously published article, we seek to understand the response mediated by the interaction of different polymorphisms at the same time, in relation to the clinical features of patients with cystic fibrosis (CF). To answer this question, we used a bioinformatics tool called MDR (Multifactor Dimensionality Reduction), which aims at the interaction of genetic and/or
environmental agents in response to clinical severity, which should be divided into two groups: (0) minor severity; (1) greater severity. By assumption, the distribution of data enables the comparison between a combination of these polymorphisms, as well as, interaction between them in the clinical response of the patient. However, the characterization of each individual polymorphism in association with clinical patients given by numeric and categorical distribution of data enables the understanding of each specific polymorphism in association with the clinic. The studies conducted to date, seek the direct association of each polymorphism with the clinic, especially lung disease mediated by pulmonary function test, given by spirometry (especially considering FEV\textsubscript{1}%). And, the analysis, in the present study, allows comparison with other studies conducted in other research centers and geographic regions. Finally, the \textit{GSTM1} and \textit{GSTT1} deletion polymorphisms, which are compared separately and in clusters, and that the \textit{GCLC} gene was not analyzed as a modifier in CF by other research until the publication of the article cited by the reviewer. In addition, the analysis of haplotype polymorphisms in \textit{GCLC} gene was considered in present review.

In their current manuscript “Polymorphisms in the glutathione pathway modulate cystic fibrosis severity: a cross-sectional study” de Lima Marson and colleagues report on the role of polymorphisms in the glutathione metabolism genes \textit{GCLC} and \textit{GSTP1} as well as large genomic deletions in \textit{GSTM1} and \textit{GSTT1} on the manifestation of cystic fibrosis disease. The authors report on 180 CF patients who have been characterized for these candidate modifier genes and the \textit{CFTR} mutation genotype. Association to 28 clinical phenotypes was tested for each of these clinical phenotypes in this cohort. The manuscript is well organized and the results are presented very clearly, summarizing the findings on each genetic variation considered for each of the clinical phenotypes studied in a separate supplementary table. The data presented by de Lima Marson and
colleagues is highly relevant for those who analyze cystic fibrosis modifiers as it shows convincingly that the GST metabolic pathway influences the CF phenotype on several levels. However, at the first glance the data is confusing as the association signals do not appear consistently in all CF subgroups and the reader does not get help in understanding this in the present manuscript. Moreover, some association results substantiate each other and the reader does not get help to realize this in the present manuscript. Hence, the findings (which are currently listed comprehensively and discussed in context with other association studies that target the same candidate genes) need to be discussed more extensively in context with each other and the population specifics. – Answer: In the present study we demonstrated the direct influence of each polymorphism and the haplotype analysis for *GSTM1* and *GSTT1* - (apart from the inclusion of haplotype analysis for *GCLC* gene), in association with the clinical variables of CF patients. In the study, the gene cluster for mutations in *CFTR* gene is necessary, which included three groups for analysis: (i) two mutations identified in *CFTR* gene belonging to Classes I, II and/or III, (ii) one identified mutation in *CFTR* gene belonging to Class I, II or III; (iii) no *CFTR* mutation identified. The groups difficult to understand the data, and often becomes complex understanding of the article. However, the division in groups is peculiar for CF disease when we analyze modifier genes, which together with environment and *CFTR* mutations characterize the disease in its different degrees of clinical severity. It is worth emphasizing about the description of the study population, which has its origin just a research center, which comprises patients in long-term monitoring, with a mixed population, and clinical importance for the study of gene modifiers. In order to facilitate understanding, better description of the data in the session data was performed, and the discussion session consider new paragraph addressing the importance of polymorphisms in general and was also made.
discussion about the admixed population analyzed as follows, in topics, subsequently throughout this review.

Major compulsory revisions:

1. The authors group their patients prior to analysis with respect to the number of resolved *CFTR* mutations. Does the frequency of pancreatic sufficient patients vary between these patient subsets? This might help to explain the findings of an association signal in one patient subgroups only based on an underlying survivor effect and/or selection bias. Example: Table 6, GCLC-129C>T / association signal with phenotype *P. aeruginosa* mucoid, observed only in the subgroup “one CFTR mutation identified”.

The reader might be puzzled: If this GCLC polymorphism modifies CF disease, why is this signal absent among those patients with two identified mutations who are typical CF patients? What makes these subgroups heterogeneous? There are currently three possible explanations:

1A. About the only phenotype that can be reliably predicted based on the *CFTR* mutation genotype is the pancreatic status, given rise to the rule that those patients who do carry one mild *CFTR* mutation (class IV and V) are pancreatic sufficient. Table 5 states that 80% of patients in this study are PI, consequently 20% are PS. This is a high prevalence of PS in comparison to the Caucasian CF populations (incidence of PS genotypes 5 - 10%) and might reflect the high incidence of mild *CFTR* mutations that cause pancreatic sufficiency in the Brazilian CF population. Alternatively, it might reflect that some patients are carriers of CF mutations who suffer from another disease but cystic fibrosis and not CF patients. These misdiagnosed individuals would then be pooled in the groups “one *CFTR* mutation identified” or “no *CFTR* mutation identified” and be absent from the group “both *CFTR* mutations identified”. Consequently, the
groups defined by the authors would be highly heterogeneous and the lack of the association signal in the group “both CFTR mutations identified” would point to the fact that this modifier has no impact on classical CF, but might play a role on atypical CF / CF-like disease and other misclassified phenotypes. – The authors agree with the issues raised by the reviewer. CF patients in the group with two identified mutations in CFTR gene have the disease considered typical and assume that, in this group, only patients with two mutations included in Class I, II and/or III was enrolled. About the other groups of patients, the possibilities raised by the reviewer are, in short, two: atypical CF, with the presence of mutation Class IV, V or VI, or the presence of diseases that cause changes of sodium and chlorine, similar to CF, and in this case, the presence of mutations in CFTR gene would not be positive and plausible for analysis. For the question (1) we agree with the reviewer, and there is really the possibility of ambiguous results regarding the determination of CFTR mutations to different degrees of severity of CF in these cases, patients in the groups "identified one mutation" or "no mutation identified" may present any mutation in CFTR gene related classes described, which enable the analysis of data differently, considering as the best group of patients, if possible, the complete genotype determination. In the case of question (2), the possibility of the patient, unless CF is reduced in our study, based on the following settings: (i) all patients had levels of sodium and chloride in sweat bigger than 60 mEq/L in two or more samples collected on different days, (ii) patients presented with CF clinical features, mainly chronic obstructive pulmonary disease, presence of bacteria in sputum, change in spirometry, comorbidities ( such as; osteoporosis, nasal polyps, diabetes mellitus, pancreatic insufficiency); (iii ) in the case of patients with one or no identified mutation in CFTR gene, was performed dosage of active CFTR via rectal biopsy, and all patients included in the study had abnormal values for the biopsy, ( iv )
in the last case, nasal potential was performed in some patients, and all values were "changed", but it points out that we have no control standard curve for analysis. Considering the presence of pancreatic insufficiency as a variable closely associated with the presence of severe mutations, actually in our study, we have a high prevalence compared to other research centers originating patients, so almost exclusively with Caucasian origin. In our population, there is a high rate of miscegenation and variation between patients, and perhaps, that is closely associated with higher frequencies of mutations Class IV, V and/or VI. To prove this fact, it would require the complete sequencing of CFTR gene, for all patients in our center, having as a basis for diagnosed by sodium and chloride. Given this analysis will be possible to show different groups of patients according to class of CFTR mutations. The study for the determination of mutations have been made, and in the future, we will have a better basis for to develop scientific studies considering the modifier genes in CF. However, in this manuscript, we bring attention to new analysis which was conducted before the review, patients with a mutation identified in CFTR gene and those with none, were grouped based on the presence of pancreatic insufficiency. In this case, it was disregarded patients who were not pancreatic insufficient, with the assumption, better standardization of data to be analyzed. Data are presented in the results and discussion session. All attached tables show the results of the new analysis.

1B. If CFTR transports GSH the finding (-129C>T / association signal with phenotype
P. aeruginosa mucoid, observed only in the subgroup “one CFTR mutation identified”) might reflect that the GSH pathway modifies CF in a subgroup of patients who display CFTR mediated residual function (class IV and V mutations on one CF chromosome, phenotype PS), presuming that these individuals are more prevalent in the group one CFTR mutation identified? Than two CFTR mutations identified?. It would also point
out that CFTR can indeed mediate GSH transport, at least in some tissues as residual
CFTR activity is likely to lead to higher GSH transport capabilities of the mutant CFTR
protein. - Answer: The transport of GSH CFTR-mediated is of utmost importance for
the cellular antioxidant response and epithelial tissue. In CF, oxidative stress affects the
lung parenchyma and subsequently epithelial cells due to the accumulation of mucus
and secretions. For patients with "residual" CFTR protein, we would have to
extracellular oxidative stress response, better performance, because as pointed out by
the reviewer, would be favorable to the passage of GSH to the outside. However, in
patients with CF with two CFTR mutations, only alternate routes would be activated for
the passing of GSH. Even though, most of GSH, is transferred to the external
environment via CFTR in cases of residual CFTR, would be little changed from the
activity of GSH, since it is known that less than, 5% occurs for expression of CFTR in
minor severe mutations class, thus, considering 65% for the GSH passing via CFTR, we
would have a percentage response to GSH presence in the external environment of at
most 3.25%. In any case, the analysis performed excluding the presence of patients with
pancreatic insufficiency provides better grouping of patients and optimizes the response
of the associations in this study. The text was made to include excerpt explaining the
role of GSH by the activity of CFTR protein.

1C. The patients were recruited cross-sectionally and hence, mild CFTR mutations
might be overrepresented in this patient panel (indicating that patients who carry a
severe CFTR mutation genotype have died with a higher likelihood than patients with
mild CFTR mutations prior to being recruited for this study). This would also explain
convincingly why the prevalence of PS patients is high in this patient panel. Again, this
would than lead to an asymmetric distribution of PS patients in the subgroups stratified
for CFTR mutation genotype. – Answer: The authors agree with the reviewer and, based
on the new approach adopted for the analysis of clinical variables, we believe it is optimized presentation of the data, as well as, the approach of this study.

The authors have access to the data to discriminate between 1A, 1B and 1C:

1A? review clinical charts for patients without identified CFTR mutations. Misdiagnosis with a sweat test value > 60 mval/l occurs, but hardly if the sweat test is taken repeatedly. uses for false-positive sweat tests are: atopic dermatitis and malnutrition and several other conditions (e.g. listed in: O'Sullivan BP, Freedman S D. Cystic fibrosis. Lancet. 2009 May 30;373(9678):1891-904. doi: 10.1016/S0140-6736(09)60327-5.) - Answer: To confirm the data presented in this study, the medical records of all patients (with one or no identified mutation in the CFTR gene) were reviewed. In all cases, addition of Sodium and Chlorine in sweat have “changed” values, all patients had symptoms suggestive of CF. To confirm the hypotheses, patients had data for active CFTR present in the gut epithelium analyzed, this being another study of our group in partnership with Margarida do Amaral and Karl Kumselmann. In this study, all patients had levels of CFTR protein negative in biopsy. Although is not a validated diagnostic test for CF, but that is enable to possibility a better data description and characterization of our sample.

1B: count prevalence of identified class IV / class V mutations in subgroups “both CFTR mutations identified” versus “only one CFTR mutation identified” versus “no CFTR mutation identified”. Likewise, assuming that some PS CFTR mutations have not been identified, count the prevalence of PS phenotype in subgroups “both CFTR mutations identified” versus “only one CFTR mutation identified” versus “no CFTR mutation identified” – Answer: The description of the patients in the study, based on the presence of pancreatic insufficiency and classification by CFTR mutations was performed and presented in this session results.
1C: rank patients according to date of birth. If PS patients are more frequently among early-born individuals, this indicates a survivor effect in the population. Likewise, if patients with two unresolved CFTR mutations and/or patients with one resolved CFTR mutation and/or patients with two identified CFTR mutations cluster by birth cohort, this indicates a survivor effect in the population. – Answer: In the study, to understand the frequency for pancreatic sufficiency, having fixed the age of the patient as parameter was calculated, the age difference between the groups of patients with and without insufficiency (Table 13). In the case of groups with and no identified CFTR mutations, after exclusion of patients with pancreatic sufficiency no difference in age was observed.

To clarify the reviewers assessment: the data is, according to the reviewer’s opinion, valid and true, i.e. there is no flaw in the analysis. However, the reviewer urges the authors to evaluate their valuable data pool comprehensively to elucidate why the association signals are observed in one patient subgroup only. This is likely best done in a separate discussion paragraph. Apart from the aforementioned finding (GCLC-129C>T / association signal with phenotype P. aeruginosa mucoid, observed only in the subgroup ?one CFTR mutation identified?), the other findings for which a similar discussion of elevated frequency of PS cystic fibrosis is necessary are: GCLC-3506A>G / P. aeruginosa no mucoid / One identified CFTR mutation GSTM1 deletion polymorphism / SpO2 / No CFTR mutations identified GSTT1 deletion polymorphism / P. aeruginosa no mucoid / No CFTR mutations identified GSTM1/GSTT1 deletion polymorphism / Bhalla score / No CFTR mutations identified and complementary for GSTP1+313A>G / Osteoporosis / Two identified CFTR mutations for which the prevalence of osteoporosis in the other two patient subgroups (One identified CFTR mutation, No CFTR mutations identified) might differ and this difference in prevalence
might explain the result. – Answer: The inquiry performed is valid and extremely important for the present and future studies. Thus, the new data presentation, the discussion was held with groups whereas patients with pancreatic insufficiency in the case of one and no CFTR mutation identified. For different groups that were addressed in the question, the following paragraph was included in the study: “The PI was used in statistical analysis as factor correction for no determination of CFTR mutation in CF groups with no or one CFTR mutation screened. After the patient exclusion to statistical analysis all the previous positive association were negative, except for GSTMI null allele. The null allele was associated as protector factor for onset of digestive symptoms (OR=0.134; CI=0.023-0.606; Table 2).”

2. The interrelationship of the association findings is not clearly discussed in the present manuscript. For instance, finding of an association for two different phenotypes that describe the same pathology substantiates such a result considerably: Two different GCLC variants are associated with two different P. aeruginosa phenotypes (albeit in the same patient subgroup “one identified CFTR mutation only”). This appears to be confirmatory of the interrelationship of GCLC and P. aeruginosa colonization in CF? Can the authors construct haplotypes of these two SNPs and deduce which haplotype one is protective against P. aeruginosa and which one promotes chronic colonization?

Answer: The analysis was performed and is presented in Table 9 for all variables. In the text, the following paragraph was added: “GCLC haplotype analysis for GCLC-129C>T and GCLC-3506A>G showed association for A. xylosoxidans and CC + AA genotypes (OR=17.9; CI95%=2.781-411.6; Table 3).”

Minor points:

3. Page 8, text reads: “adjusted by the Bonferroni correction (# corrected = 0.05/number of tests”)? the tables imply that alpha observed was divided by 4 to display the corrected
P-value? Please also state the rationale for the number 4 (as some readers might expect correction for 28 clinical variables, or worse, for 4 X 28 tests). – Resposta: No parágrafo foi adicionado o trecho: “O valor de α foi corrigido levando em consideração a análise dos mesmo dado no mesmo grupo de pacientes mais que uma vez, sendo: (i) grupo de todos pacientes; (ii) pacientes com nenhuma mutação identificada no gene CFTR; (iii) pacientes com uma mutação identificada no gene CFTR; (iv) pacientes com duas mutações identificadas.” - Answer: In this study no correction was made for 28 clinical variables because the main question is whether the polymorphisms result in action in CF severity for different markers, and in spite of clinical markers, are mainly of chronic obstructive pulmonary disease, not there is a relationship of direct dependence. However, when directly comparing the group of patients divided by CFTR mutations, we only do association to a variable at the same time, but for four groups, so the correction factor is important in this context.

4. Page 8, text reads “Data distribution showing a high standard deviation was analyzed by the median value.” This is not interpretable - was the primary data transformed? If so, how? Please explain differently. - Answer: Statistical analysis of the data there was no transformation of the same, except when there was a high standard deviation, the data were considered in two groups divided by its median. Thus, the following paragraph was added: “Data distribution showing a high standard deviation was analyzed in groups distributed according by the median value.”

5. page 10, text reads “The COPD pathophysiology is similar to CF in that it involves cellular responses, inflammatory mediators, and oxidative stress” There are clearly many observations that point to differences in COPD and CF. Please reword to avoid that readers who do not have a profound knowledge of the pathophysiology of CF and COPD (such as pure geneticists with limited clinical expertise) can confuse these two
different diseases in the future. - Answer: To avoid errors in the interpretation, the paragraph was rewrite for: “The COPD pathophysiology is similar, in some aspects, to CF in that it involves cellular responses, inflammatory mediators, and oxidative stress”.

6. page 10, text reads: “little is known about GST genes and CF severity?”. “Little” is unspecific - of course, the GST genes deserve a more thorough look as a modifier, but the analysis of nearly 2000 CF patients for the role of the candidate modifier genes GSTM1 does not qualify as “little”. Please reword. – Answer: The paragraph was rewrite to: “However, there is no mention in the scientific literature of GCLC polymorphisms as clinical modulators of CF severity, and is necessary new studies to illuminate about GST genes and CF severity.”

7. page 11: text reads “previously investigated in 1,940 children (aged 8-11 years).” Please quote the population & reference 37 already here. – Answer: The reference was altered in text position. Also: as this is a very large study, do the results presented within this manuscript identify agree with this previously published work in an entirely different study population with respect to the assignment of the risk allele at GSTM1, GSTT1 and GSTP1?

8. page 11 and following: PAM and PAMN are defined in the manuscript, but cryptic to read within the text. Please avoid these abbreviations and use full-text instead. - Answer: In the text were replaced acronyms to facilitate understanding.

9. page 11, text reads “higher frequency of the PAM to CC genotype” Please specify the SNP. – Answer: The paragraph was rewrite to: “In the present study, it was associated with a higher frequency of the PAM to CC genotype for GCLC-129C>T polymorphism in patients with one CFTR mutation identified (Table 2; p=0.044).”

10. page 12, text reads ?the GSTP1+313A>G polymorphism was associated with a low risk of osteoporosis (p=0.036; with two CFTR mutations identified) as a protective
factor and with young age \( \leq 154 \) months \( (p=0.044; \text{without taking the CFTR gene into account}) \) as a risk factor. This needs to be edited? the A allele is protective against osteoporosis? The A allele is increased among young patients with unresolved CFTR mutation genotype? – Answer: The paragraph was rewrite to: “In this context, in our data, the A allele is protective against osteoporosis, and is increased among young patients with unresolved \textit{CFTR} mutation genotype.”

11. page 12, text reads “The presence of osteoporosis is influenced by several different factors, including mutations in the \textit{CFTR} gene, the environment, modifier genes, and increased life expectancy [39].” Please add a comment here with respect to the observations of the osteoporosis incidence in your patient subgroups (see above, comment 1: “\textit{GSTP1}+313A>\textit{G} / Osteoporosis / Two identified \textit{CFTR} mutations for which the prevalence of osteoporosis in the other two patient subgroups (One identified \textit{CFTR} mutation, No \textit{CFTR} mutations identified) might differ and this difference in prevalence might explain the result.” Any survivor bias in this population will emphasize the effect of a modifier on a condition associated with longevity. – Answer: In the present study, we forward the amendment containing the frequency of osteoporosis in different groups and the paragraph: “The present study found that the AA genotype of the \textit{GSTP1}+313A>\textit{G} polymorphism was associated with a low risk of osteoporosis \( (p=0.036; \text{with two \textit{CFTR} mutations identified}) \) as a protective factor and with young age \( \leq 154 \) months \( (p=0.044; \text{without taking the \textit{CFTR} gene into account}) \) as a risk factor. The G allele, however, is responsible for increased \textit{GSTP1} expression. The presence of osteoporosis is influenced by several different factors, including mutations in the \textit{CFTR} gene, the environment, modifier genes, and increased life expectancy [46]. In this context, in our data, the A allele is protective against osteoporosis, and is
increased among young patients with unresolved CFTR mutation genotype. The osteoporosis frequency is shown in table 12 and the Table 3.

12. page 12, text reads: "recent discoveries indicate that CFTR modulates the transport of GSH, creating a dysfunction in the antioxidant defense." Please give references to substantiate this claim. Level of interest: An article of importance in its field – Answer: The reference was used: Gould NS, Min E, Martin RJ, Day BJ: CFTR is the primary known apical glutathione transporter involved in cigarette smoke-induced adaptive responses in the lung. Free Radic Biol Med. 2012; 52(7):1201-1206.”

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests: I declare that I have no competing interests.

Reviewer: Andreas Hector

Reviewer's report:

This study by Fernando AL Marson and colleagues describes associations of SNPs in the GSH pathway with clinical parameters of CF patients. Various SNPs were associated with e.g. PSA infection, the Bhalla score, oxygen saturation or osteoporosis. Yet, in this reviewer's view, some aspects are lacking in this manuscript.

Major Compulsory Revisions:
Although the authors use the Shwachman-Kulczycki score and other clinical scores, in this reviewer's opinion, lung function tests would be preferable for this kind of studies. Large gene modifier studies in the USA use preferentially longitudinal lung function data as described by Schluchter et al. (AJRCCM 2006). This reviewer would also rather choose this clinical parameter (for example estimated FEV1 at age of 20 years) for genetic association studies instead of cross-sectional analyses. Surely, longitudinal lung function data are not available in all CF centers, but the authors should at least discuss why they have not taken the lung function tests into account in this study. And if longitudinal lung function is not available, association with cross-sectional lung function should be included. – Answer: In this study were used as markers of pulmonary function in association with the different polymorphisms analyzed the following variables: transcutaneous oxygen saturation and spirometry test data (FEV1%, FVC%, FEF25-75% and FEV1/FVC). All spirometry data were used based on the predicted value based on the norms published by Polgar. For the polymorphisms analyzed, considering our population, there was no association with spirometry. In the study, spirometry test was performed in a cross sectional study. The longitudinal study, really, as reported by the reviewer, is the best parameter for data analysis associated with the severity of pulmonary manifestations of CF patients in the long term. However, considering the reality of our center, we cannot perform this analysis for the present study. However, current studies of our research group intend to analyze the pulmonary function test every year, and in the future it will be possible to perform the analysis described. In this context the passages were added: “Spirometry was performed in patients older than seven years of age with the CPFS/D spirometer (MedGraphics, Saint Paul, MN, USA) and data were recorded using the PF BREEZE software version 3.8B for Windows 95/98/NT [33]. The following variables were included: forced vital
capacity [FVC (%)]; forced expiratory volume in the first second [FEV1(%)], the ratio between FEV1 and FVC(%) [FEV1/FVC(%)]; and forced expiratory flow between 25 and 75% of the FVC [FEF25-75%]. The data was analyzed considered international curves values for spirometry tests [34,35].

Study limitations: (i) CFTR mutation with no complete screening; (ii) short population of CF patients; (iii) spirometry test performed by transversal method and did no performed longitudinally; (iv) no measure of GSH activity or GST and GCLC proteins, taking into account the sample collection limitation in our center and time to process all data. Study highlights the data by: (i) one CF center collection – considering an admixed population, the CF patients from one center minimizes miscegenation factors. Another fact, is the similar environmental and the same access to treatment; (ii) the high number of clinical markers evaluated provides better association and characterization of modifier genes action; (iii) complete CF diagnosis performed by different methods.

The authors conducted statistical analyses by subdividing the CF population by the numbers of CFTR mutations they have found in each patient (0, 1 or 2 mutations identified). Unfortunately, to this reviewer it is not clear, why the authors chose this method. Why would the numbers of found mutations be important for the effect of gene modifier? Please explain in the manuscript. – Answer: CF severity is affected by three main factors, namely the environment, modifier genes and CFTR mutations. CFTR Mutations are grouped into six classes. Classes I, II and III show absence of CFTR protein active in the epithelium, being associated with more severe clinical pulmonary disease, worst progression of the patients, increased frequency of comorbidities and the presence of higher risk for pancreatic insufficiency. At the same time, Classes IV, V and VI exhibit residual CFTR protein and can ameliorate the clinical disease. In this context, the statistical analysis approach four groups: (i) the population of patients with CF
excluding the CFTR mutations – without taking into account CFTR gene, (ii) patients with no identified CFTR mutation, (iii) patients with one CFTR mutation identified, (iv) patients with two CFTR mutations identified.

Furthermore, are there any associations between the CFTR mutations and the SNPs? Is the OR of the various SNPs higher (or lower) for example in dF508 carrying individuals (homozygous/heterozygous) compared to non-dF508 subjects? – Answer: There was no association, and p-values are shown in table 1.

One limitation of this study is the low numbers of CF patients included. This should be discussed by the authors.

Minor Essential Revisions:
There was a formatting problem with the legend of Table 1 - it was hidden behind the table. – Answer: The table was revised and adjusted.

Discretionary Revisions:
In gene modifier studies, it is also very important to investigate functional effects of the SNPs. Because all the SNPs assessed in this study are involved in the GSH pathway, it would be very helpful to measure GSH/GSSG levels in blood or even airway samples of the CF patients. Alternatively or additionally activity of the enzymes could be measured. This would most likely further increase the quality of this study. – Answer: The authors agree with the possibility of investigating the effects on functional level by the SNPs. However, being a chronic disease, is difficult to approach and interaction of research and treatment, to collect material for functional studies in the blood, and especially airways becomes complicated. Since the DNA is used as a “Bank” from Laboratory of Medical Genetics, upon approval of the parent or guardian, after collection and analysis of principal CFTR mutations is easy to determine the SNPs. In some cases a second test is performed, where it would be possible for functional
analysis by collecting more material, but the degree of patient acceptance is not high.
Moreover, the collection of other routine tests, which could be done to analyze the
material for functional study, is not integrated into sector of genetics on hospital. For
future studies, different proteins being analyzed expression indirectly by
polymorphisms will be analyzed and the study will address the functional analysis.

- The authors describe the association between PSA (non-mucoid and mucoid) with
different SNPs. However, was there a difference for different infection stage (for
example according to the consensus paper by Doering et al. (JCF 2012) - chronic
infection > 50 % positive microbiological analyses in the last 12 month, intermittent
infection > 0 and # 50 %, negative or never)? - Answer: We follow the consensus and
the section was added to the article: “Several clinical variables were employed,
including Shwachman-Kulczycki, Kanga and Bhalla clinical scores [31]; body mass
index (BMI) [for patients older than 19 years, the BMI = weight/(height)2 formula was
used, while remaining patients used the WHO ANTHRO program (children 0–5 years
of age) or the WHO ANTHRO PLUS program (children 5–19 years of age)]; patient’s
age (≤154 and >154 months); time to diagnosis (≤24 and >24 months); time of first
clinical symptoms (digestive: ≤3 and >3 months; pulmonary: ≤6 and >6 months); time
to first colonization by Pseudomonas aeruginosa (≤31 and >31 months); bacteria in the
respiratory airways: mucoid P. aeruginosa and no mucoid P. aeruginosa, Achromobacter
xylosoxidans, Burkholderia cepacia and Staphylococcus aureus - the positive status was
evaluated considered chronic infection (patients in whom more than 50% of the
preceding 12 months was culture positive) + intermittent infection (patients with less
than 50% of cultures positive ). A patient was negative considering as free of bacterium
(when no bacterium was grown from samples in the previous 12 months, despite a
history of prior colonization) + never infected (patients in whom the bacterium) has
never been cultured, i.e. this consensus was formulated for P. aeruginosa, but in our data was used for all bacteria [32]; transcutaneous hemoglobin oxygen saturation (SpO2) and spirometry variables.”

In this reviewer's opinion, the study by Gu et al. (Nature 2009), which is an important recent gene modifier studies in CF, should be discussed in this manuscript, too. The latter study was a whole genome study and, to this reviewer's knowledge, none of the investigated genes in this manuscript was found to be relevant for the CF lung disease. – 


Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests: I declare that I have no competing interests.

Reviewer: Harriet Corvol

Reviewer's report:

Major Compulsory Revisions
As highlighted in this study, it is now recognized that, besides the CFTR genotype, other genetic factors could contribute to the cystic fibrosis (CF) phenotype. The authors studied whether genes associated with the glutathione (GSH) metabolism could influence CF severity. While this study may be of interest, major commentaries could be raised:

- Compared to the current published studies on modifier genes in CF, the cohort is very small: 180 patients and no replication? it is well known that such a small cohort could provide false-positive results, especially if there is no replication in an independent population. There are actually large CF cohorts that are being internationally collected with a main aim to study such modifiers of the CF disease phenotype. The authors should at least acknowledge such collections and cite proper publications (and, please, do not cite reviews but the original papers)!  

**Answer:** The authors acknowledge the collaboration and believe to be relevant exception made. CF population studies have expanded the understanding of the disease and enabled a greater understanding of the variables associated with the severity of the disease, particularly by modulating of modifier genes. The studies are mostly international multicenter allowing increase in the number of patients analyzed. In this study, a population was analyzed and characterized by variable clinical taking into account the modulating by gene modifier genes and **CFTR** mutations. Considering the environment, we also have the conditioning variable, however, considering "only" a reference center, we have a homogeneous approach with regard to access to treatment of the disease and less variation in the external environment. The same cannot be said of large-scale studies, where treatment is appropriate, but associated with each center in a unique way, in addition, we must consider the genetic variation of agents polymorphisms between the populations when considered over. We need to improve the data in our study, and in this case, is important
highlights the number of clinical variables analyzed – 28 – and the same cannot be addressed in other studies. Finally, in our study, was statistically correcting for the analyzes performed, minimizing the effect of “small” population and the sample calculation, taking into account - minor allele frequency – was performed, what is consistent with the possibility of carrying out the study, and the calculations for statistical analysis.

- For example, a North-American GWAS in a very large CF population did not show any association with lung disease severity in the regions studied by these authors (Wright and coll. Nat Genet 2012) ? underlying again the risk of false-positive associations driven by such a small cohort without replication? - Answer: The study was not considered in the discussion, making reservations to employee working model in which the analysis was - Presence of \textit{P. aeruginosa} in CF versus gene variants by exoma, being technically differentiated the method used and the questions asked.


- A lot of different phenotypes have been studied: far too many in this far too small cohort. One might expect such a broad phenotype's study only in a much larger cohort. - Answer: The CF severity is variable, and the search for markers of severity needs to be broad in order to understand the complex phenotypic expression of the disease, which may be mediated by modifier genes. By restricting the clinical analysis, we can "lose" viable markers for genotypic comparison, included in this study. The clinical
characterization is one of the most relevant point in CF disease and the association with it taking into account \textit{CFTR} mutations and polymorphisms in modifier genes is important.

- The authors have corrected their results for multiple testing using a Bonferroni correction, but what is the ?number of tests? referring for? - Answer: The number of tests is four, considering the independence between the clinical variables analyzed, and based distribution groups for the different polymorphisms in the genes analyzed as modifiers and possible genotypes for the \textit{CFTR} gene.

- The authors studied CF patients with and without pancreatic insufficiency, but it is well known that severity of the CF patients is highly linked to this clinical feature and, at least, a correction should be made? - Answer: The correction was performed on all study and changes are described in the text and supplemental tables.

- It is unclear how did the authors chose the SNPs within the genes studied? - Answer: The choice of polymorphisms was performed based on their characterization in the population in different studies, considering respiratory diseases and other diseases with significant oxidant response.

- What are the equation used to calculate the percent predicted FEV1 and FVC? Why do not use, as frequently used nowadays in CF gene modifiers papers, FEV1 CF-specific percentiles? – Answer: To calculate the values of spirometry test was used to standardize of the ATS and ERS (2013) for the technical approach, and data analysis considered the percentage of predicted values, referring to the population curve of Polgar.

Level of interest: An article of insufficient interest to warrant publication in a scientific/medical journal
Quality of written English: Not suitable for publication unless extensively edited. – The article was corrected by Edanz Groups.

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests: I declare that I have no competing interest