Sefydliad Cenedlaethol ar gyfer Ymchwil Gofal Cymdeithasol ac Iechyd | National Institute for Social Care and Health Research

Llywodraeth Cymru
Welsh Government

National Institute for Social Care and Health Research (NISCHR)

Applicant's Response to Peer Review Comments

Project reference: **HF-14-29**

Lead applicant: **Arwel Wyn Jones**

Project title: **Novel biomarkers and frequency of exacerbations in chronic obstructive pulmonary disease**

Applicant's response to peer reviewers
Project reference: HF-14-29
Lead applicant: Arwel Wyn Jones

Comments:

I (the applicant) and the academic mentors would like to thank the reviewers for the careful evaluation of our initial submission and providing us with specific comments and suggestions to improve the suitability of the application. Please find the responses below which have been prepared to address all of the reviewers' comments under each assessment criteria. For purposes of clarification, reviewers' comments are in bold font. We would like to thank the reviewers for the positive comments on our previous submission and we now firmly believe that the outlined amendments have further strengthened a project that will contribute to a deeper understanding of the underlying mechanisms responsible for susceptibility to exacerbations in COPD.

A: Scientific merit

Peer review 1

Clearly a lot of thought and effort has gone into working up this application and so I don't want to seem like I'm picking holes. Although the aims are clearly stated, in essence this is essentially a bit of a fishing trip and I guess if you measure enough markers something is bound to pull out. Now that's not meant to be a damming criticism because we've all been there and done that, but I don't believe this will lead to any real impact on terms of personalised medicine per se. Other groups are doing similar studies creating databases of local and systemic biomarkers and then linking to health informatics, and in much larger numbers than the present proposal. There are no ethical concerns. I have my doubts within the constraints of the limited sample size that this is going to achieve high impact in for example the Blue Journal.

Response

We thank the reviewer for acknowledging that the proposal has clearly defined aims. The key issues under 'scientific merit' relate to i) data analytic approach and ii) placing the proposed research into the context of ongoing research.

i) Non-targeted, or unbiased, metabolomics enables reliable identification of small molecule biomarkers characteristic for a particular physiological state in response to internal or external perturbations. Although increased understanding of COPD has been due to some hypothesis testing, numerous recent and ongoing studies have undertaken an unbiased exploration (metabolomics) to determine the extent to which COPD exacerbations represent clinically meaningful subgroups/phenotypes. Our data analytic approach achieves maximum coverage and elucidation of the metabolome's chemical space, and is the state-of-the-art method in this field.

ii) We do not agree that other groups are doing similar studies to ours, but with much larger sample sizes. We are aware of two current multi-centre, non-interventional, longitudinal studies that have recruited/are aiming to recruit larger cohorts of patients than our present proposal, e.g. ECLIPSE N = 2164; SPIROMICS N = 2400). These larger studies have, however, been designed to address primary aims and objectives that differ from those of the proposed study. For the investigation of novel biomarkers (i.e. primary aim of the proposed study, but secondary objectives in the above studies), smaller, nested participant samples are being used: ECLIPSE n = 201 and SPRIOMICS n = 300.

The scientific merit of this proposal is reflected by the uncertainty in the clinically important area of exacerbation risk in COPD, balanced against the enormous potential of metabolomics to improve phenotypic characterisation for better diagnosis, risk assessment and efficacious preventive intervention.

Peer review 2

The aims of the project are clearly stated and the study seems well-designed although I find the timeplan too ambitious. The research addresses an important gap in our knowledge and could provide very useful insight. I see no ethical problems and the research is of interest in the UK as well as outside.

Nevertheless, I think there is one important shortcoming in the project – and this should be addressed. However, I think the study proposal can with little effort be changed to address my concern. I am mainly worried that the outcome, COPD exacerbations, are too heterogeneous events to be lumped together. It seems highly likely that different biomarkers will predict different types of exacerbations and that different pathways are involved. I think the project should address the heterogeneity of exacerbations by e.g. trying to type them by presence/absence of viral infection, bacterial infection and other treatment as this in combination with slightly better phenotyping of patients beyond frequent/infrequent exacerbators is necessary. The phenotyping could include type of previous exacerbations, treatment of these, +/- colonisation (clinically) in stable state and underlying treatment.

We thank the reviewer for recognising that this is a well-designed project that has clear aims to address an important gap in our knowledge of COPD and provide insightful findings to users beyond the UK. The reviewer had one primary concern regarding the phenotyping of exacerbations.

We accept the suggestion that better phenotyping of exacerbations may be necessary, and we will adopt patient diary cards for recording of exacerbations prospectively and use Anthonisen criteria to classify exacerbations as Type 1, 2 or 3. (Anthonisen et al., 1987; Wedzicha et al., 2013)

We agree that COPD exacerbations can be heterogeneous events and that different biomarkers and pathways may be involved. However, whilst different types of exacerbation may have distinct biomarkers, these will most likely have limited potential for proactive, preventive intervention, since they are event-specific and,

almost certainly, proximal to exacerbation.

The greatest potential for proactive, cost-effective preventive intervention for COPD exacerbations would derive from identification of biomarkers of more distal common provenance. Indeed, frequent exacerbators are characterised not by a predisposition to one type of exacerbation over another, but by increased risk of exacerbation per se, i.e. independent of type. Thus, the overarching aim of the proposed study is to investigate potential differences between exacerbator types (frequent and infrequent) rather than types of exacerbation.

As there is a lack of robust evidence to guide decision-making in this context, we will undertake hypothesis-generating analyses of potential biomarkers for sub-groups of participants stratified by exacerbation type. In these analyses we will focus on confidence intervals to assess the likely range of effects, rather than the significance of p values, and interpret the findings with due caution.

.

B. Methodology

Peer review 1

The methodology is well described but the power calculations are in my opinion rather optimistic and dare I say it a bit vague. I say this because there is no mention a priori of any primary outcome variable on which the study has been powered, in terms of the minimal important difference and the associated SD. Also I think the authors should consider over-powering at 90% given this is a small single centre study .The reason I say this is because it is highly likely that the power will depend not only the standardised response mean ratio of the different variables being measured, but also on the heterogeneity of the population being studied in terms of GOLD stage. I see no attempt to stratify for GOLD stage and inevitably we will be left with weakened conclusions based on post hoc subgroup analysis according to GOLD stage –again I'm speaking from personal experience here from this type of study analysis . Hence the two groups inevitably will not be matched with the more frequent exacerbators being in GOLD 3/4, which in turn has inherent problems in terms of potential confounders such as comorbidity and therapy –eg ICS and infection risk .For example performing 16S sputum microbiome analysis and neutrohphil function is likely to be heavily influenced by ICS therapy. A much larger sample size will be required to tease out these type of factors in the analysis. My steer would be to recommend they greatly increase the sample size by bringing on board other centres in Wales who I'm sure would be happy to collaborate in this type of national study which is relatively easy to do as it does not involve a CTIMP . Although they are stratifying according to exacerbations, I can see no mention of any measurement of health status ,QOL or functional status (e.g. 6MW or shuttle test) being measured here . Also I would suggest they also measure effort independent pulmonary function –such as RVC , RV/TLC ,R5 ,AX which might be more sensitive than conventional tests such as FEV1 and FVC . Perhaps the authors should also think about including some other relevant biomarkers such as Gallectin, HS-CRP and BNP– it would be shame not to at least bank some blood for analysis later on.

We thank the reviewer for recognising that the methodology is well described. The reviewer raised some key issues which were related to (i) power calculations ii) stratification iii) sample size iv) measurement of other parameters.

i. We are pleased to be able to clarify our power calculations. The project aims to identify, or discover, potential differences in immune parameters between frequent and non-frequent COPD exacerbators. There is, therefore, no *a priori* primary outcome, since identifying novel candidate biomarkers is the purpose of our investigation. We do not consider it necessary to 'over-power' the study (90% rather than the conventional 80%), since it will not be a single centre study (as noted by the reviewer) and, more to the point, to do so would be inconsistent with its exploratory purpose and would offer no clear benefit – see Table 1 below.

Table 1. Sample size and standardised effect size calculations.

Power	**Effect Size**				
	0.4	**0.45**	**0.5**	**0.55**	**0.6**
80%	200	160	128	106	88
90%	263	210	168	140	116

ii. We agree that there is much heterogeneity across GOLD stages 1 to 4, but do not consider it necessary to stratify by GOLD stage. As our eligibility criteria specify, only patients with GOLD stage 3 (severe) or 4 (very severe) will be eligible for enrolment in the study. Excluding GOLD stage 1 (mild) and 2 (moderate) necessarily reduce heterogeneity in the study population. We will, however, control for GOLD stage in the analyses.

iii. We think that it would be premature to greatly increase the sample size by bringing on board other centres in Wales'. Our previous research has identified mean differences in biomarkers of immunity between frequent and infrequent exacerbators equivalent to standardized effect sizes of 0.58. We have powered the study to detect differences around this magnitude of effect – see Table 1. More specifically, to detect a standardised effect size of 0.5, as statistically significant at $p<.05$, requires 128 and 168 participants with 80% and 90% power, respectively.

As detailed in the response to reviewer 2 in section B, we anticipate being able to recruit between 200 and 240 participants during the six month recruitment period. Allowing for 15% loss to follow-up, we will be able to detect effect sizes of at least 0.45 with 80% power, or 0.5 or more with 90% power, which is sufficient given the effect sizes we have

previously observed in a comparable sample. We will, in contrast to the original submission analyse candidate biomarkers (neutrophil function, sputum microbiome, vaccination responses) in all patients rather than a sub-group. Logistically, this will not create any further issues as the initial application did aim to collect blood and sputa samples at these timepoints (i.e. metabolomics).

iv. Given our goal of developing the area of COPD research in Wales, as mentioned in our original submission, blood and sputa will be banked for later analysis to stimulate future research or ancillary projects that may measure other relevant biomarkers (e.g. CRP). Although we recognise the differences in measures of disease severity (e.g. effort independent pulmonary function, FVC), we do not believe that this modification would be consistent with our intention to validate biomarkers of disease activity and hence do not consider the additional time and training required to both incorporate these additional measures and ensure optimal technique within the project to be worthwhile.

Peer review 2

Methodology is well described, ambitious and the applicant / team seem capable of taking this forward.

We thank the reviewer for the positive comments.

C. Planning and resources

Peer review 1

The funding request seems reasonable for what has been proposed.

We thank the reviewer for their comment.

Peer review 2

The project seems well planned although the timetable seems ambitious given the number of subjects that the applicant wants to recruit. I therefore have slight worries as to the time plan. The budget seems very reasonable but may need slight expansion if the project has problem with time plan. The project would in my view offer good value for money.

We thank the reviewer for highlighting that the project is well planned and will offer good value for money. We agree that the programme of work is ambitious but organisation and structures already developed as part of earlier pilot work/collaborations mean the proposal is challenging but achievable. We will seek adoption for study onto the Welsh Portfolio and have confirmed through discussions with network research nurses that the level of support required is feasible.

We provide additional information on participant recruitment so as to demonstrate feasibility of our strategy (Table 2). Since the initial submission of the application we have identified and agreed collaborations with three other hospitals in Wales (Aberystwyth, Carmarthen, Withybush) alongside Prince Phillip Hospital. Our recruitment pool is now based on the respiratory outpatient clinics and databases of these hospitals rather than GP practices in the Llanelli area within the original submission.

We firmly believe that the information outlined here as well as support from lead clinicians at four hospital sites demonstrate the planning and resources in place for this project. Although the main approach to patient recruitment will be via routine clinical contact at hospital sites, if necessary, mailouts will be sent to potential participants from consultant registers.

Table 2. Participant recruitment

	Recruitment Period (months)					
	1	**2**	**3**	**4**	**5**	**6**
Sites recruiting	1	2	3	3	2	1
Monthly n	20	40	60	60	40	20
Cumulative N	20	60	120	180	220	240

The key features of our recruitment strategy are as follows:

- 4 sites will recruit for 3 months each, with one site launched per month for 4 months.
- Target recruitment is 20 participants per site per month.
- There is built-in contingency to prevent having to extend the recruitment period, with sites 1, 2 and 3 continuing recruiting for the duration of the recruitment period if necessary, i.e. if n <100 at 3 months.

The changes in the number of patients for measurement of candidate biomarkers outlined in Section B (Methodology, Peer reviewer 1) and inclusion of other hospital sites here (Planning and Resources, Peer reviewer 2) has resulted in the following changes in costs of the proposal:

- Cost of consumable has increased by £24,624.54.
- Travel costs have decreased by £500.00 due to the lower total mileage to collect samples at four hospital sites at each 3 month timepoint compared to Llanelli alone.
- Total cost of the application NISCHR has increased to £281,664.15 from £257539.91 (9% increase).

D. Impact and dissemination

Peer review 1

I have serious doubts that the data generated from a relatively small dataset in a heterogeneous population are going to result in important advances in health or related social care services. The plans for dissemination seem appropriate. A priori stratification according to GOLD stage with proper power calculations would also be helpful in terms of getting the results published in a high impact journal.

We thank the reviewer for reporting that our plans for dissemination are appropriate. The key issues raised by the reviewer were related to impact: i) concerns about the small heterogeneous dataset, ii) advances in health and social care, and iii) publication.

i. We have responded to the issue of sample size in other sections (Section B - Methodology) but we would again emphasise that our proposed sample is in line with norm of the field whereby a systematic review of 93 biomarker discovery studies reported a mean sample size of 204 (range 40-500) (Hundt et al., 2007). It is also worthy to note that sub-studies nested within larger longitudinal studies of COPD patients (as mentioned in section A) consisted of a sample with greater heterogeneity (GOLD stage II-IV rather than only III and IV) than proposed here.

ii. We are aware that exploratory analyses may not lead to immediate impact on health or related social care services but do emphasise that these projects lay the foundations for more applied, clinical and service organisation and delivery research. We can provide an example of such advances by work carried out (Metabolic Diagnosis of LUNG, MEDLUNG) from the research team (Mur Lab) on identifying relevant pathways in respiratory disease which have been based on smaller sample sizes (25 patients, 25 controls) than the present proposal. This employed metabolomics (Fourier Transform Infrared spectroscopy, FT-IR) to optimally discriminate between sputa of cancer and non-cancer 'healthy' control cases (Lewis et al., 2010). The protocol has been patented and its exploitation as a viable screen is being explored with commercial and clinical stakeholders.

iii. We have addressed the issues of prior stratification and power calculations in an earlier section (Section B - Methodology) and firmly believe that the gap in research knowledge being addressed by this project would be highly publishable and encourage further research as noted by reviewer 2.

Peer review 2

The study – if carried out according to plan – is very likely to have an impact and will stimulate further research. However, this is based on the assumption that further subtyping of exacerbation events is possible (as outlined in A). With studies like this there will always be some concern about how to implement results in future strategy but it does not seem impossible and it seems outside the scope of the current project to actually address this in detail.
The plan for result dissemination seems appropriate.

We thank the reviewer for the positive comments on the potential impact of this project. We have previously accepted the consideration of further subtyping in section A. We would also like to reiterate the reviewer's point that this project will stimulate further research whereby health care professionals can make more informed decisions regarding patient management and care pathways.

-

E. Overall Impression

Peer review 1

I honestly do not believe this is going to make a real difference to clinical care in terms of health outcomes. My advice would be to make this a Welsh collaborative study with an appropriate level funding and much bigger patient numbers. Then you would have something with significant potential national and international impact, rather than a small data trawling exercise as it currently stands.

We have dealt with responses to larger sample size and impact in previous sections (Methodology, Impact and Dissemination). The project strongly follows calls for research efforts to be adjusted to take into account disease activity as well as disease severity. Further refinement is required of clinically important phenotypes through identification of biomarkers that act as surrogate endpoints and/or establishment of candidate biomarkers to serve as intermediates in intervention studies. We propose the validation of candidate systematic and pulmonary biomarkers that have been shown by the research team and others to have strong biological plausibility in terms of their role in the pathogenesis of COPD and/or are biologically relevant and sensitive to inform an important gap in research relating to frequency of exacerbations (host defence and infection risk). This approach coupled with expertise of the research team in exploiting the promising technology of metabolomics, clinical practice and public health engagement will accelerate the translation of knowledge from the basic sciences to its application in clinical and community settings. We have outlined changes to the sample frame (Section B - Methodology) (from GPs to Consultant clinics) for both efficiency and wider representation of the target Welsh population. Rather than being a small data trawling exercise based on convenience samples, this project will collect quality data using repeated measurements in a well-characterised population, and employ state-of-the-art data analytic techniques.

Peer review 2

Strengths: Well picked topic, good eye on current weakness and good methodology, adequate supervision.
Weakness: Ambitious / over-ambitious timeplan and concerns about lack of exacerbation phenotyping.

We thank the reviewer for acknowledging the need for this project and recognising the capability of the team (methodology and supervision) in delivering the research objectives. The reviewer raised two principle concerns: feasibility/recruitment and COPD phenotyping.

i) We have widened our recruitment pool and provided additional information regarding participant recruitment, particularly, the feasibility of achieving the target sample size including mitigating strategies to ensure recruitment is achieved within the specified time (Section B).

ii) We have addressed the concern with phenotyping within the study design (Section B) by incorporating Anthonisen criteria to distinguish different types of exacerbation, and through hypothesis-generating analyses of subgroups of participants stratified by exacerbation type.

References

Anthonisen, N. R,. Manfreda, J., Warren, C. P., Hershfield, E. S., Harding, G. K. and Nelson, N. A. (1987). Antibiotic therapy in exacerbations of chronic obstructive pulmonary disease. *Annals of Internal Medicine*, **106**, 196–204.

Hundt, S., Haug, U. and Brenner, H. (2007). Blood markers for early detection of colorectal cancer: a systematic review. *Cancer Epidemiology Biomarkers and Prevention*, **16**, 1935-53.

Lewis, P. D., Lewis, K, E., Ghosal, R., Bayliss, S., Lloyd, A. J., Wills, J. and Mur, L. A. J. (2010). Evaluation of FTIR Spectroscopy as a diagnostic tool for lung cancer using sputum. BMC Cancer, **10**, 640.

Wedzicha, J. A., Rabe, K. F., Martinez, F.J., Bredenbröker, D., Brose, M., Goehring, U. M. and Calverley, P. M. (2013). Efficacy of roflumilast in the COPD frequent exacerbator phenotype. Chest, **143**, 1302-11.