CONCEPT PAPER TEMPLATE

Provisional Paper Title:	Cluster analyses of respiratory function and symptoms
Proposing Author:	Tony (Xian) Zhang
Author's Phone, Fax or E-mail:	xian0727@gmail.com
Date:	29/6/2017
P.I. Sponsor (if the proposing author is a student or colleague of an original PI)	Bob Hancox

Objective of the study:
To identify phenotypes of lung function abnormalities and respiratory symptoms at different ages in adulthood and to assess the stability of these over the life course to age 45.
To further assess predictive/risk factors for each phenotype at each age and to model transitions between these phenotypes if they are not largely stable.

Data analysis methods:
1. Cluster analysis (initially using hierarchical clustering) of measures of lung function (FEV_1, FVC, etc.) and symptoms (wheeze, cough, sputum) at each adult age in the Dunedin study.
2. Assess the stability of the phenotypes represented by these clusters over time (i.e., the extent to which Study members move between clusters).
3. Identify predictors (e.g. smoking, asthma, sex, BMI, atopy) for each cluster at each age using multinomial logistic regression.
4. Identify predictors of transitions between clusters, if numbers permit, using state-transition models.

Variables needed at which ages:
Lung function – all ages
Basic demographic and height and weight data
Atopy
Asthma diagnoses and respiratory symptoms
Smoking, cannabis use.
Inflammatory markers, blood eosinophils[1]

Significance of the study (for theory, research methods or clinical practice):

There is a great deal of interest in identifying different phenotypes of common respiratory diseases. It is believed that diseases such as Asthma and COPD may not be single entities, but umbrella terms for overlapping "endotypes". In the drive for more personalised medicine, there is interest in identifying the underlying diseases to better target treatment. Cluster analysis has been used to try to separate these diseases into different subtypes.[1] However, recent evidence suggests that the reproducibility of these subtypes is poor between cohort.[2] Even the basic distinction between asthma and COPD can be difficult to make in practice, leading to the introduction of the term "ACOS" for the overlap syndrome. Even at age 38, the Dunedin study found that many participants with childhood asthma met criteria for COPD diagnoses.[3]

Most cluster analyses have been done in disease-specific subgroups (e.g. participants with COPD diagnoses). It is not clear how much these participants overlap with people without these diagnoses. It is also unclear how consistent the clusters of disease phenotypes are over time. This is the aim of this study.

If suitable stable clusters are identified, the next step will be to describe the characteristics of these and assess risk factors for cluster membership—such as smoking history and other exposures. This information will be valuable for future attempts to distinguish subtypes of respiratory disease and personalise medicine.

References:

1. Weatherall M, Travers J, Shirtcliffe PM, et al. Distinct clinical phenotypes of airways disease defined by cluster analysis. *Eur Respir J* 2009; **34**(4): 812-8.
2. Castaldi PJ, Benet M, Petersen H, et al. Do COPD subtypes really exist? COPD heterogeneity and clustering in 10 independent cohorts. *Thorax* 2017.
3. Hancox RJ, Gray AR, Poulton R, Sears MR. The Effect of Cigarette Smoking on Lung Function in Young Adults with Asthma. *Am J Respir Crit Care Med* 2016; **194**(3): 276-84.