Linking US registry data with geomarker data: longitudinal associations between environmental factors and lung function decline

European Cystic Fibrosis Conference

Eleni-Rosalina Andrinopoulou, Pedro Afonso, Rhonda Szczesniak, Grace C. Zhou, John P. Clancy, Anushka Palipana, Erika Rasnick, Cole Brokamp, Patrick Ryan, Ruth Keogh

10 June, 2022

www.erandrinopoulou.com eandrinopoulou@erasmusmc.nl @ERandrinopoulou

Introduction: Motivation

A lot of information is available

➔ Electronic medical records

Introduction: Motivation

A lot of information is available

➔ Electronic medical records

Different types of information

- ➔ Baseline characteristics: Sex, F508del, SESlow, Enzymes
- ➔ Biomarkers: FEV_1 % pred
- ➔ Nutritional status: BMI percentile
- ➔ Geomarkers (environmental/community factors): Deprivation index

This research is supported by the National Institutes of Health / National Heart, Lung and Blood Institute (grant R01 HL141286)

www.erandrinopoulou.com eandrinopoulou@erasmusmc.nl @ERandrinopoulou

Introduction: Research question

How can we appropriately connect registry data with geomarker data

Introduction: Descriptive statistics

Deprivation index (*https://geomarker.io/dep_index/*)

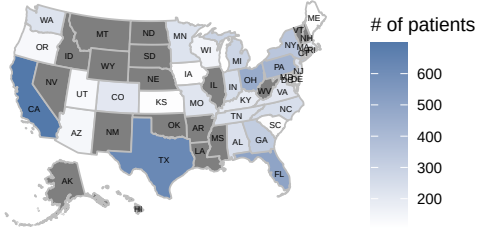
➔ Socioeconomic variables from the American Community Survey (ACS): capture "community deprivation"

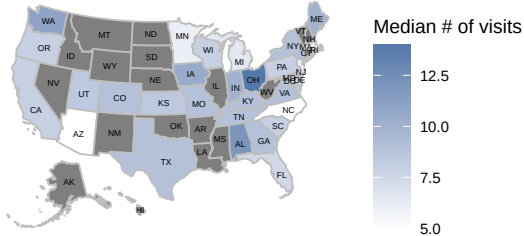
- ◇ Principal components analysis of six different 2015 ACS measures
- ◇ "Deprivation Index": the first component explains over 60% of the total variance
- ◇ Rescaling and normalizing forces the index to range from 0 to 1, with a higher index being more deprived

Cole Brokamp, Andrew F. Beck, Neera K. Goyal, Patrick Ryan, James M. Greenberg, Eric S. Hall. Material Community Deprivation and Hospital Utilization During the First Year of Life: An Urban Population-Based Cohort Study. Annals of Epidemiology. 30. 37-43. 2019

Introduction: Descriptive statistics

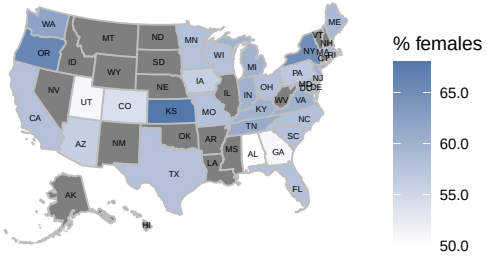
Baseline information

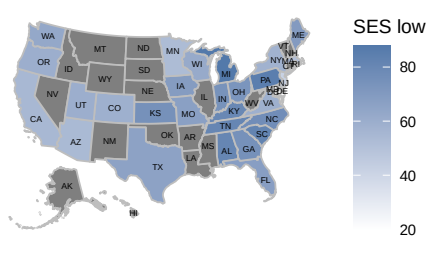




Introduction: Descriptive statistics

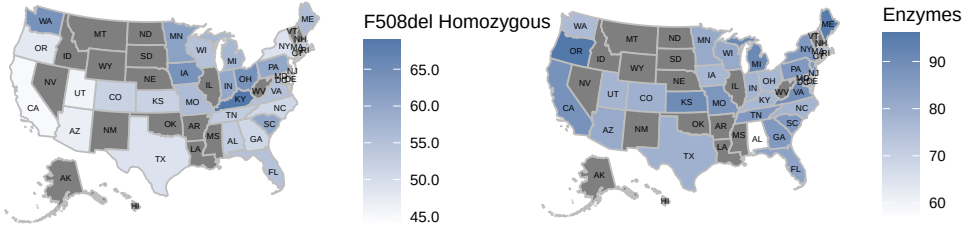
Baseline information





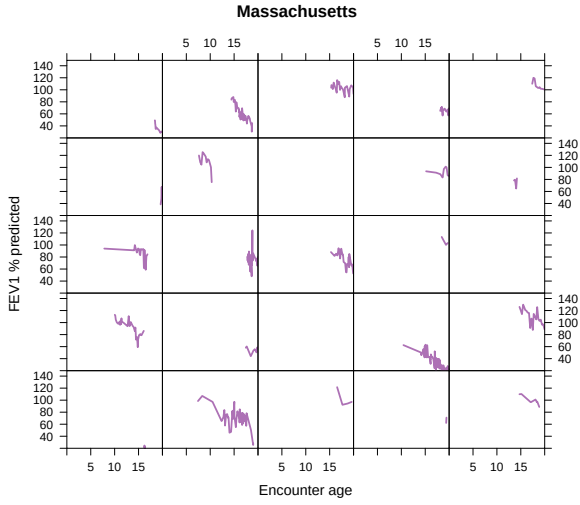
Introduction: Descriptive statistics

Baseline information



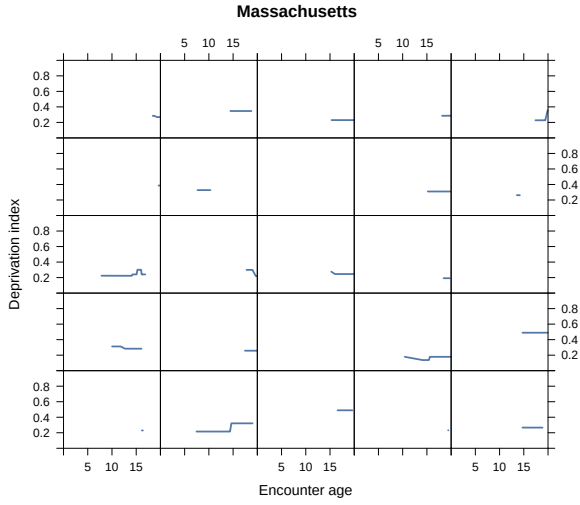
Introduction: Descriptive statistics

Biomarker



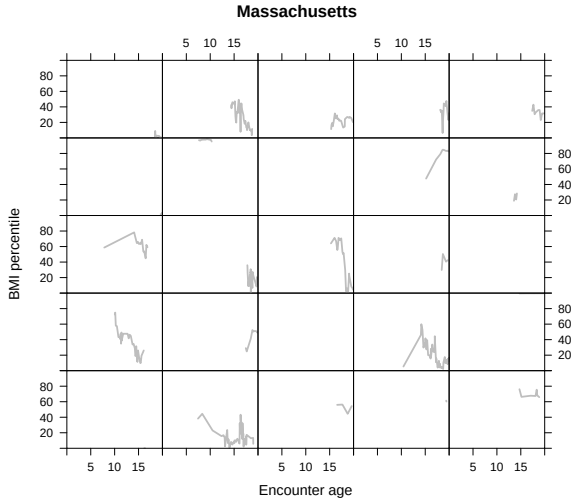
Introduction: Descriptive statistics

Geomarker



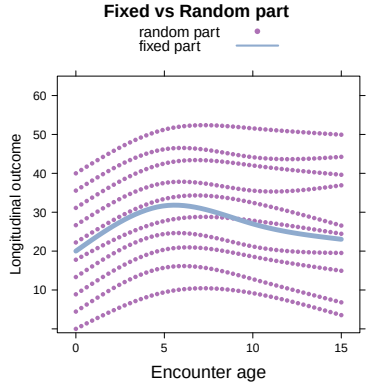
Introduction: Descriptive statistics

Nutritional status

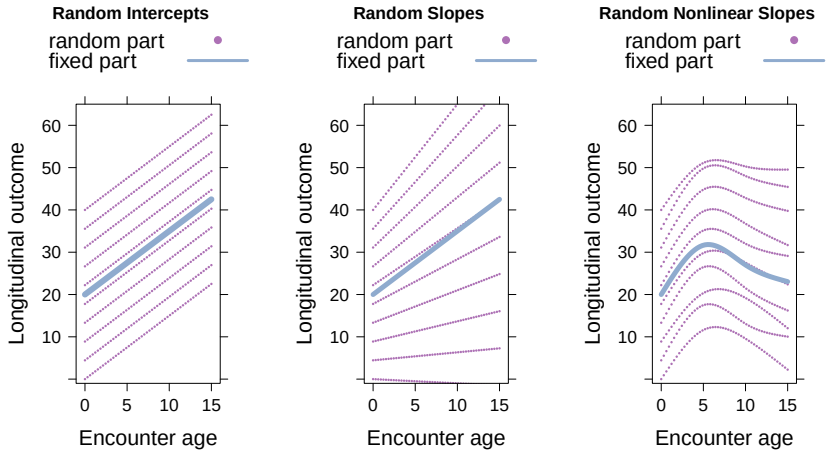


Methods: Univariate Mixed Model

$$y_i(t) = \underbrace{x_i^\top(t)\beta_1}_{\text{fixed effects part}} + \underbrace{z_i(t)^\top b_i}_{\text{random effects part}} + \epsilon_i$$



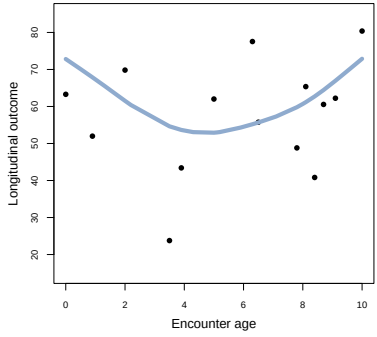
Methods: Univariate Mixed Model



Methods: Univariate Mixed Model

$$y_i(t) = \underbrace{x_i^\top(t)\boldsymbol{\beta}_1}_{\text{fixed effects part}} + \underbrace{z_i(t)^\top \boldsymbol{b}_i}_{\text{random effects part}} + \epsilon_i$$

$$= \boldsymbol{m_i(t)} + \epsilon_i$$



Methods: Multivariate Mixed Models

$$\texttt{DepIndex} = \texttt{encounterage}$$

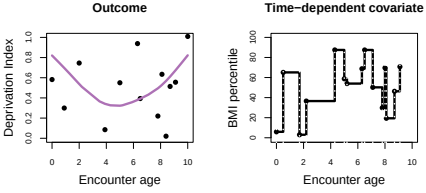
$$\texttt{FEV1\%pred} = \texttt{encounterage} + \texttt{sex} + \texttt{SESlow} +$$
$$\texttt{F508del} + \underbrace{\texttt{BMIpercentile}} + \texttt{Enzymes}$$

Methods: Multivariate Mixed Models

$$\texttt{DepIndex} = \texttt{encounterage}$$

$$\texttt{FEV1\%pred} = \texttt{encounterage} + \texttt{sex} + \texttt{SESlow} +$$

$$\texttt{F508del} + \underbrace{\texttt{BMIpercentile}}_{\text{Time-dependent covariate}} + \texttt{Enzymes}$$



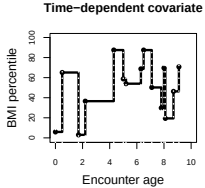
www.erandrinopoulou.com eandrinopoulou@erasmusmc.nl @ERandrinopoulou

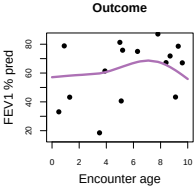
Methods: Multivariate Mixed Models

$$\texttt{DepIndex} = \texttt{encounterage}$$

$$\texttt{FEV1\%pred} = \texttt{encounterage} + \texttt{sex} + \texttt{SESlow} +$$

$$\texttt{F508del} + \underbrace{\texttt{BMIpercentile}}_{\text{Time-dependent covariate}} + \texttt{Enzymes} + \texttt{DepIndex}$$





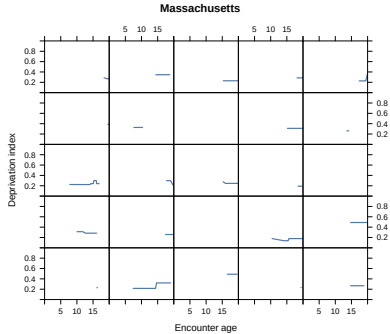
Methods: Multivariate Mixed Models

$$\texttt{DepIndex:} y_{2i}(t) = m_{2i}(t) + \epsilon_{2i}(t) = x_{2i}^{\top}(t)\beta_1 + z_{2i}(t)^{\top} b_{2i} + \epsilon_{2i}(t)$$

$$\texttt{FEV1\%pred:} y_{1i}(t) = x_{1i}^{\top}(t)\beta_1 + z_{1i}(t)^{\top} b_{1i} + \boldsymbol{\alpha}_{S2}\, m_{2i}(t) + \epsilon_{1i}(t)$$

where

- $b_i^{\top} = (b_{1i}^{\top}, b_{2i}^{\top}) \sim N(0, D)$



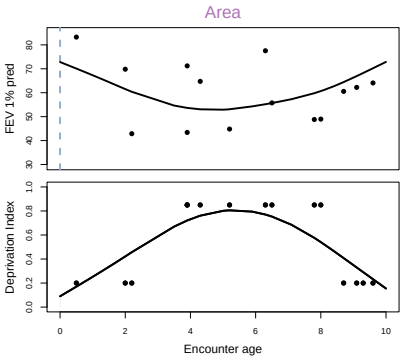
Methods: Multivariate Mixed Models

$$\texttt{DepIndex:} y_{2i}(t) = m_{2i}(t) + \epsilon_{2i}(t) = x_{2i}^{\top}(t)\beta_1 + z_{2i}(t)^{\top} b_{2i} + \epsilon_{2i}(t)$$

$$\texttt{FEV1\%pred:} y_{1i}(t) = x_{1i}^{\top}(t)\beta_1 + z_{1i}(t)^{\top} b_{1i} + \alpha_{S2} \int_0^t m_{2i}(s)ds + \epsilon_{1i}(t)$$

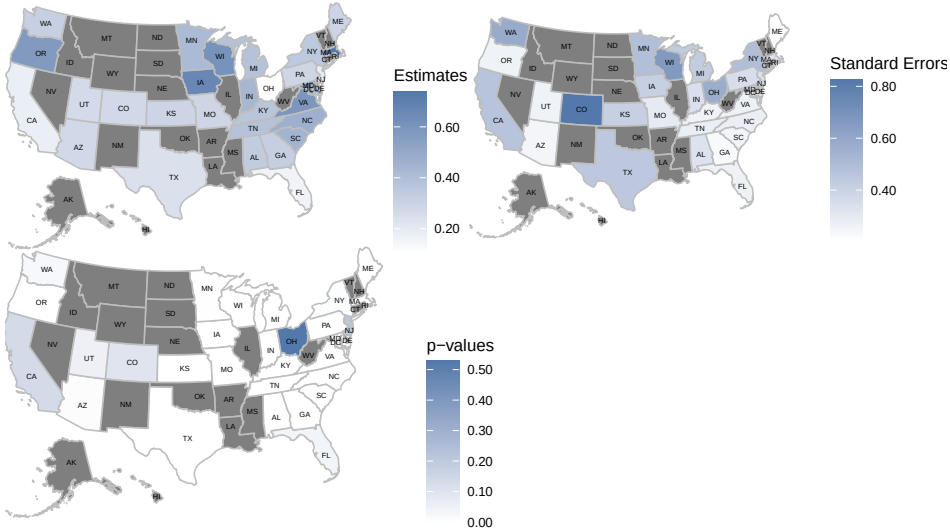
where

- $b_i^{\top} = (b_{1i}^{\top}, b_{2i}^{\top}) \sim N(0, D)$



Results: Multivariate Mixed Models

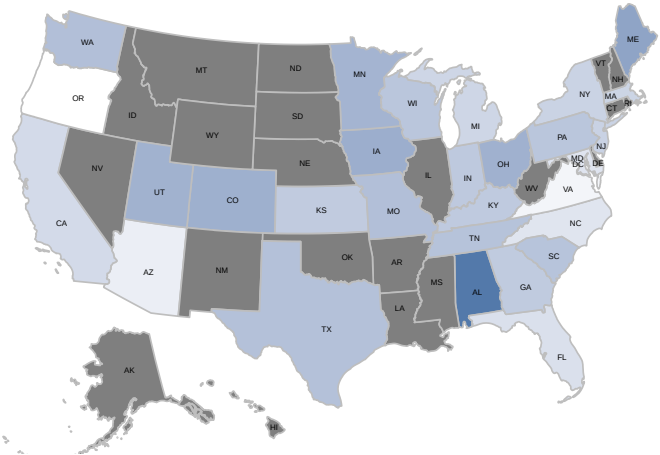
Estimate: for 0.1 unit increase in the area under the curve of the deprivation index



www.erandrinopoulou.com eandrinopoulou@erasmusmc.nl @ERandrinopoulou

Introduction: Descriptive statistics

Baseline information



Methods: Multivariate Mixed Models

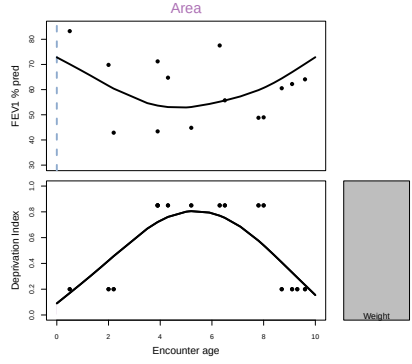
DepIndex: $y_{2i}(t) = m_{2i}(t) + \epsilon_{2i}(t) = x_{2i}^{\top}(t)\beta_1 + z_{2i}(t)^{\top} b_{2i} + \epsilon_{2i}(t)$

FEV1%pred: $y_{1i}(t) = x_{1i}^{\top}(t)\beta_1 + z_{1i}(t)^{\top} b_{1i} + \boldsymbol{\alpha}_{S2} \frac{1}{t} \int_0^t m_{2i}(s)ds + \epsilon_{1i}(t)$

Weight

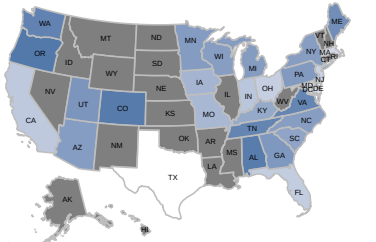
where

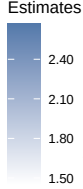
- $b_i^{\top} = (b_{1i}^{\top}, b_{2i}^{\top}) \sim N(0, D)$

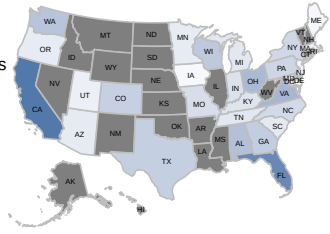


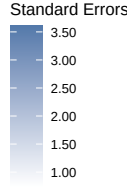
Results: Multivariate Mixed Models

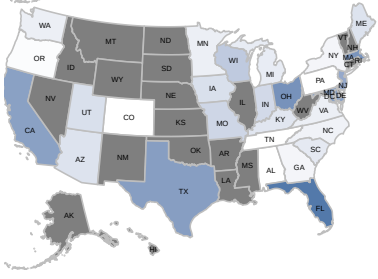
Estimate: for 0.1 unit increase in the normalized area under the curve of deprivation index

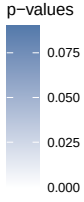












Methods: Multivariate Mixed Models

$$\texttt{FEV1\%pred} = \texttt{encounterage}$$

$$\texttt{DepIndex} = \texttt{encounterage} + \texttt{FEV1\%pred}$$

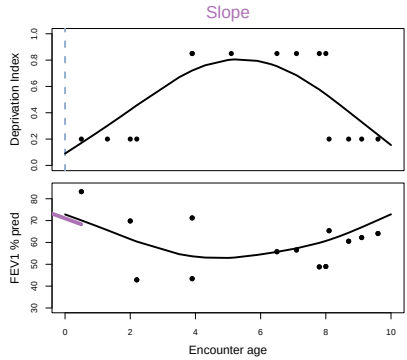
Methods: Multivariate Mixed Models

$$\texttt{FEV1\%pred:}\; y_{1i}(t) = m_{1i}(t) + \epsilon_{1i} = x_{1i}^{\top}(t)\beta_1 + z_{1i}(t)^{\top} b_{1i} + \epsilon_{1i}(t)$$

$$\texttt{DepIndex:}\; y_{2i}(t) = m_{2i}(t) + \epsilon_{2i} = x_{2i}^{\top}(t)\beta_1 + z_{2i}(t)^{\top} b_{2i} + \boldsymbol{\alpha}_{S1} \frac{d}{dt} m_{1i}(t) + \epsilon_{2i}(t)$$

where

- $b_i^{\top} = (b_{1i}^{\top}, b_{2i}^{\top}) \sim N(0, D)$



Discussion

- ➔ A lot of data is available
- ➔ Better treatment and monitoring strategies if all information is used
- ➔ Challenge in combining different types of information
- ➔ Investigate other geomarker data

Erasmus MC
University Medical Center Rotterdam

Thank you for your attention!

www.erandrinopoulou.com eandrinopoulou@erasmusmc.nl @ERandrinopoulou