PMI RESEARCH & DEVELOPMENT

Modeling Early Initiation Processes In Smoking-Induced Lung Adenocarcinomas

Vuillaume G[1], Mueller T[2], Talikka M[1], Cheng Y[1], Diehl S[2], Han W[1], Peitsch M[1], Hoeng J[1], Tobin F[3]

[1]Philip Morris International R&D, Philip Morris Products S.A., Neuchâtel, Switzerland
[2]Philip Morris International R&D, Philip Morris Research Laboratories GmbH, Cologne, Germany
[3]Tobin Consulting LLC, Newtown Square, Pennsylvania, US

Outline

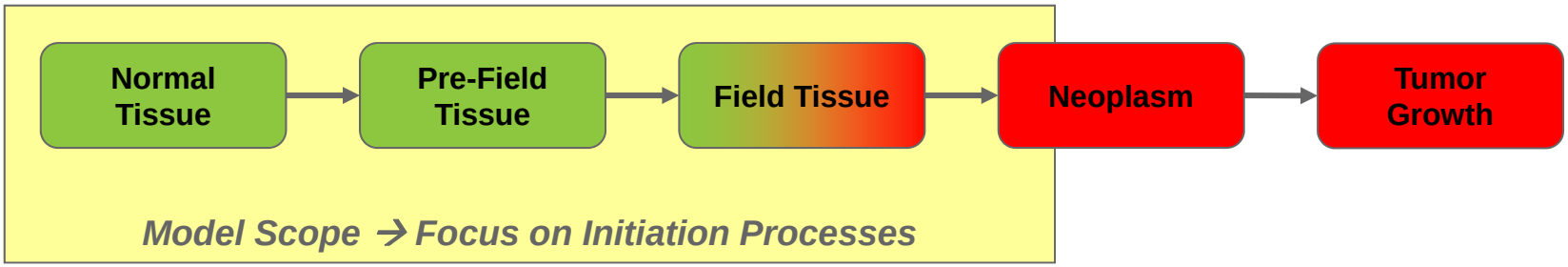
- Background
- Method
- Preliminary Results
- Summary

Cigarette Smoking (CS)-Dependent Lung Tumorigenesis Is An Extremely Complex And Poorly Understood Process

- Smoking induces histological changes…

Normal tissue

Tissue affected by CS

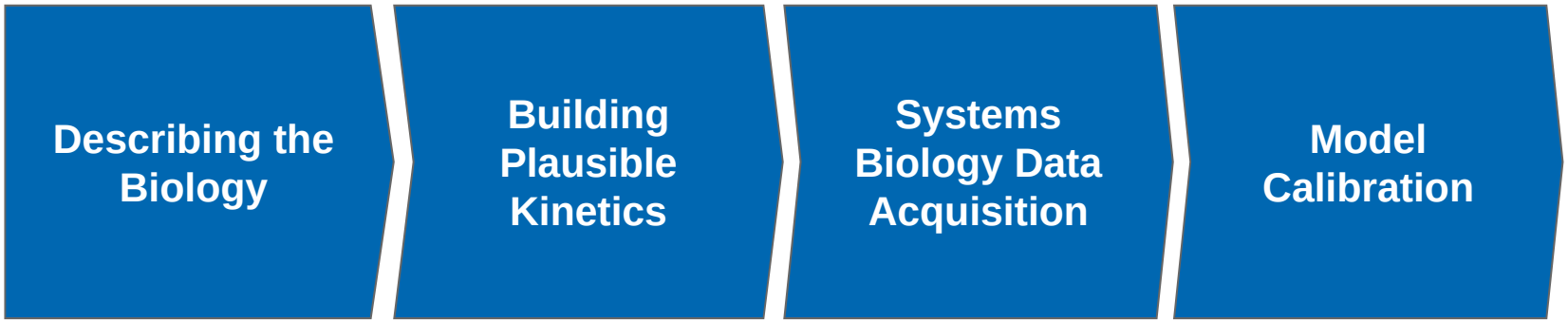


Objective

Build a model that helps to better understand the initiation processes of human lung adenocarcinoma due to smoking and to predict the time for appearance of the first neoplasm

Method Overview

- Approach consists in 4 major steps…



Description of the main biological steps leading to the disease

Translation of the biology into equations

Data acquisition to quantify and "validate" the model

Using the data to estimate the parameters of the model

PMI RESEARCH & DEVELOPMENT

Describing the Biology (1/4)

- Description of the main actors & processes of disease initiation
- Formal biological definition of the model concepts (i.e. variables)
- Design driver → *Decisions on model granularity*
  - Balance details & complexity in resulting model

Granularity in …

a) … biological entities

b) … processes

c) … in time considerations (smoking is daily and adenocarcinoma takes decades to form)

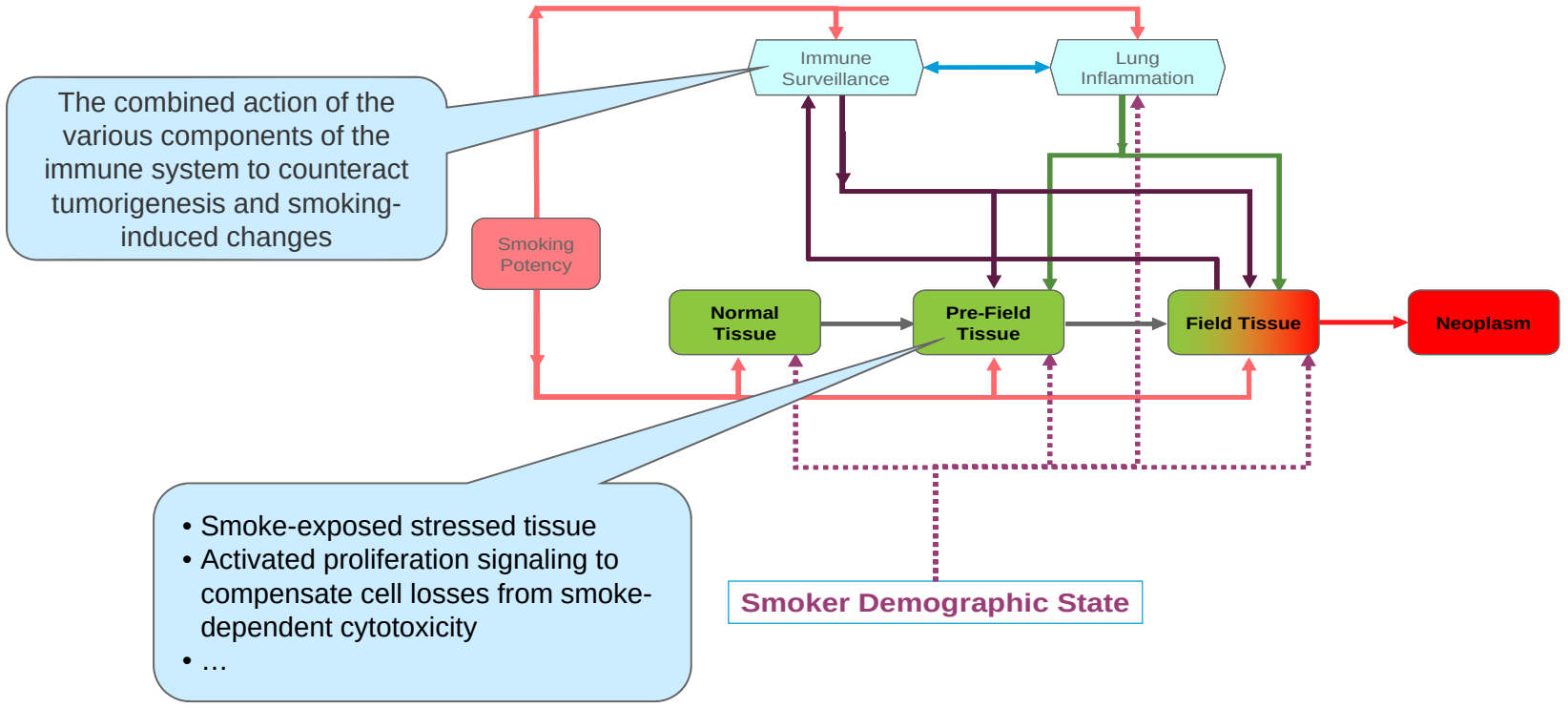
Consequences on data requirements

Abstraction & generalization are required

PMI RESEARCH & DEVELOPMENT

Describing the Biology (2/4)
First Level of Detail

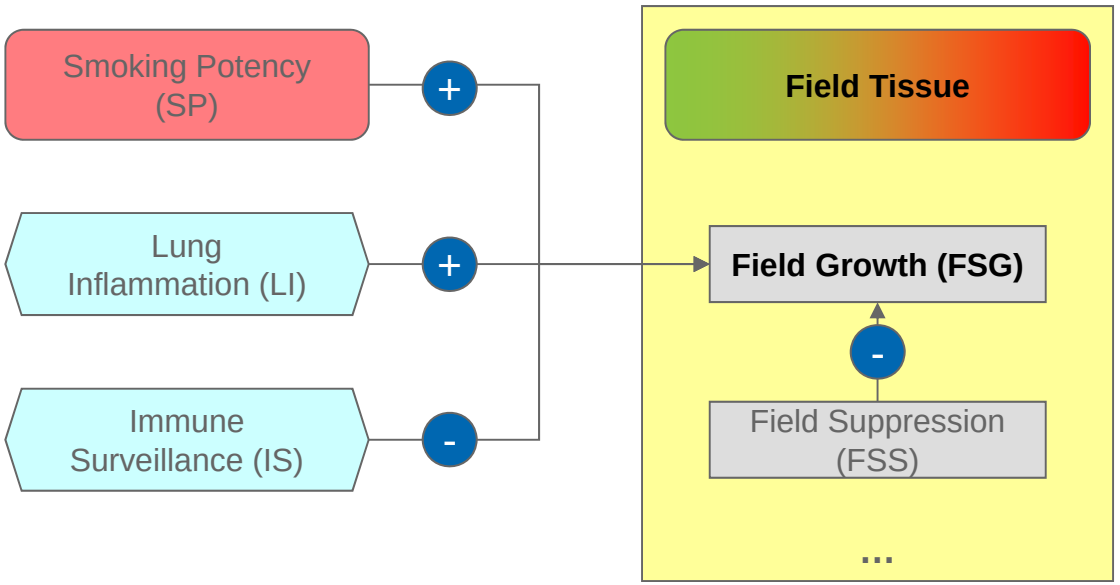
- Representation of main tissues and lung properties
- Representation of smoking (including cessation) and demographic influences



Describing the Biology (3/4)

Simple Example of Next Level of Detail

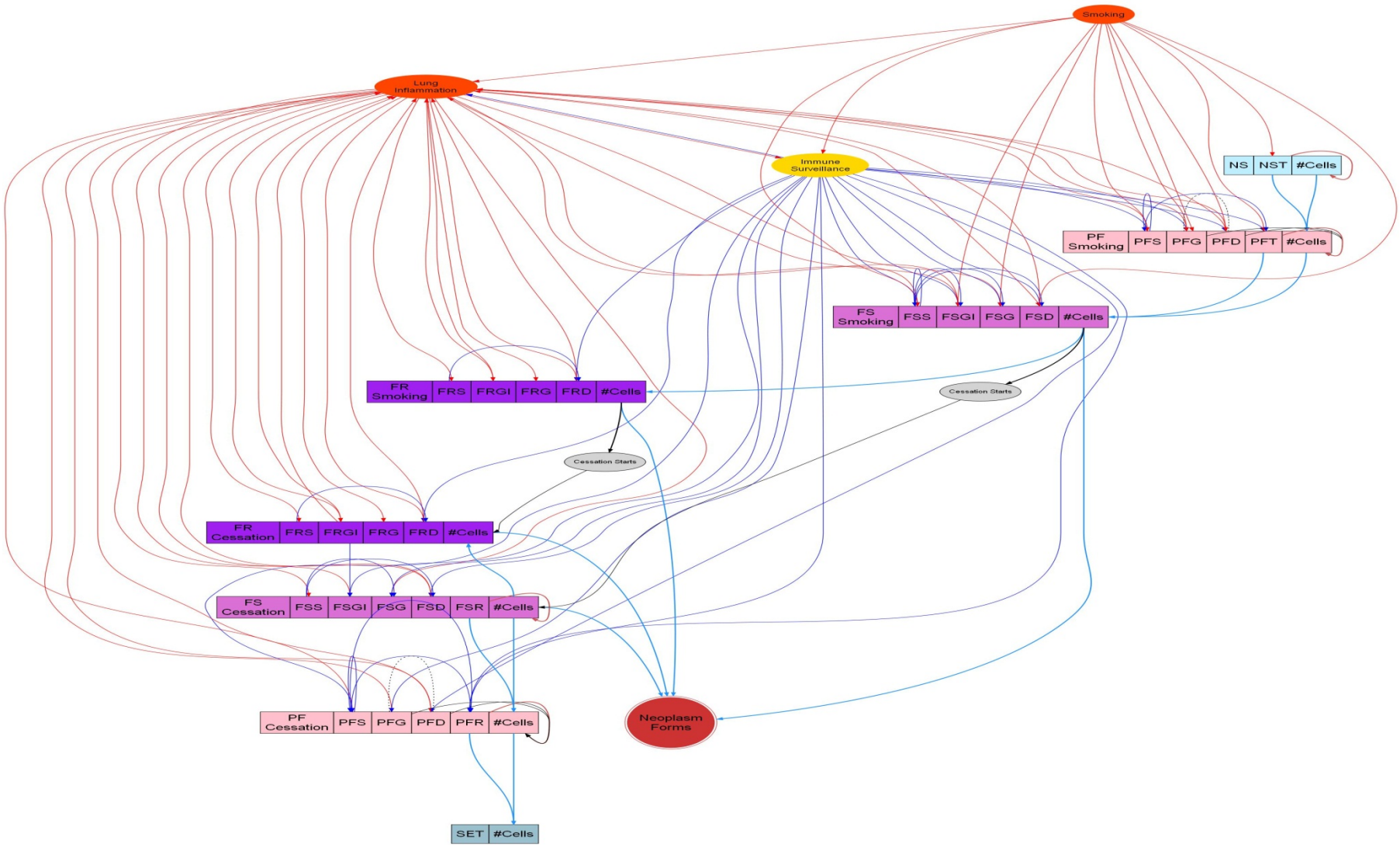
- Addition of pre-field and field tissue main properties and their interactions
- e.g. field growth and all influences on it…



PMI RESEARCH & DEVELOPMENT

Describing the Biology (4/4)

Current Model (including healing processes during cessation)



Building Plausible Kinetics (1/2)

Differential Equations (ODEs)

- Describe the quantitative evolution of the model concepts (i.e. the “variables”)
- Obtained by a “one-to-one” mapping of the influence diagrams
- Combination of first and second order effects (second order effects are of the type cell count * rate)

Neoplastic Tests

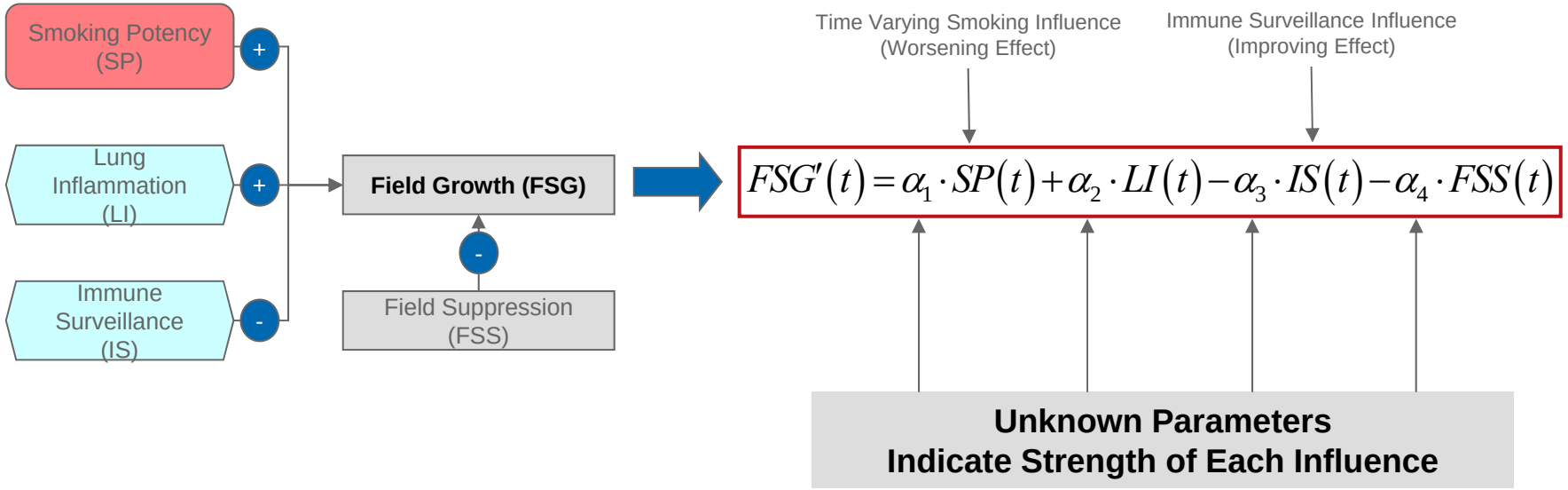
- Conditions for a cell to be considered as neoplastic
- Characteristics are translated into mathematical conditions
- *e.g. sustained growth, lack of resistance to tumorigenesis*

Constraints

- Model results have to be biologically realistic
- *e.g. maintain pre-field growth homeostasis*
- No unacceptable outputs must occur
- There are many such constraints which are derived and translated into mathematical terms

Building Plausible Kinetics (2/2)

- Example of “one-to-one” mapping…



Model contains 21 differential equations involving 59 parameters

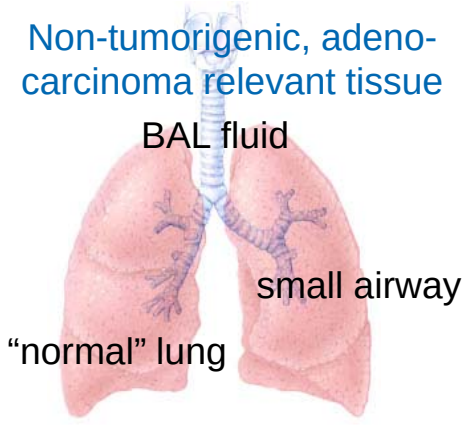
Systems Biology Data Acquisition (1/2)

- Data are required to calibrate and validate each piece of the model over all times and doses - each tissue, each property, each process

- Explicit set of criteria for data inclusion
  - E.g. focus on non-tumorigenic AC relevant tissue, quantitative smoking information, ...

- Experimental data serve as modeling surrogates
  - Diverse sources: genomics, proteomics, clinical data, pre-clinical data (include scaling to human), …
  - Biological measurements are quantitatively mapped to a corresponding model concept (e.g. IL-8 = inflammation)

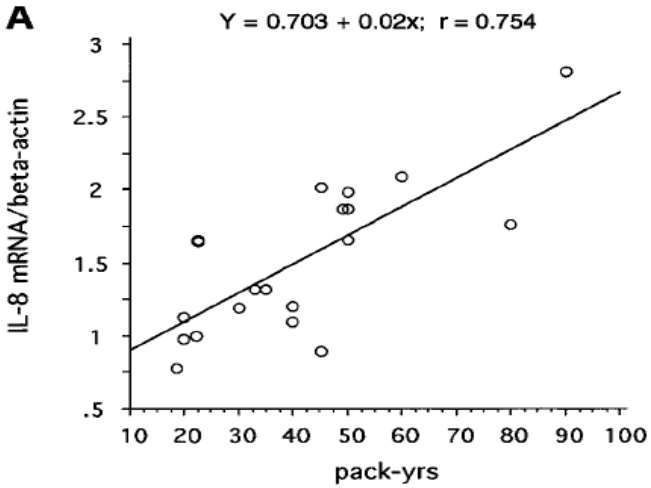
Non-tumorigenic, adeno-carcinoma relevant tissue



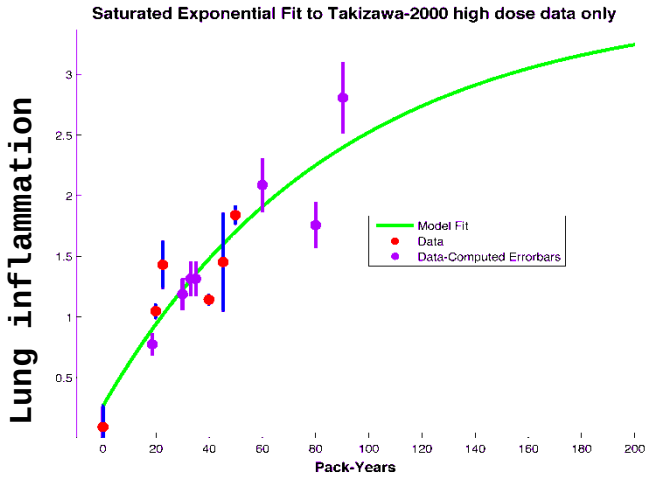
Picture taken from www.TopNews.in

Systems Biology Data Acquisition (2/2)

- Example of data mapping to model concept (sometimes an abstraction)…

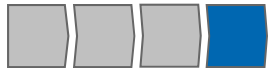


Takizawa et al., *Am J Physiol Lung Cell Mol Physiol* 278: L906–L913, 2000



PMI RESEARCH & DEVELOPMENT

Model Calibration (1/2)



- Optimization is used to calibrate the model
  - ***Find one set of parameter values that best fits each and every data set simultaneously***
  - "Single, central mechanism results"
  - All the data sets are matched to each model concept simultaneously
- Optimization constraints
  - Only plausible biological results
  - No unacceptable outputs occur
- "Validation" → provides information on where the model works well and where it does not

Model Calibration (2/2)

- Objective function of the optimization is…

$$\min_{\vec{\alpha}} O(\vec{\alpha}) = \sum_j \omega_j \cdot O\left(\left\|\text{Model\_version}(j,\vec{\alpha}) - \text{data\_set}(j)\right\|\right)$$

- … under the following plausibility constraints…

$$\vec{g}(\vec{\alpha}) = 0 \text{ and } \vec{h}(\vec{\alpha}) > 0$$

- … where O is a metric ("distance") between the model outputs and the data
- … ω_j is a weight assigned to each data set
- … and g and h are functions that ensure that biological plausibility is satisfied and that no unacceptable output occurs
  - *e.g. pre-field growth homeostasis → Pre-Field Growth (PFG) = Pre-Field Death (PFD) → PFG'(t) = PFD'(t) → constraint on the equation parameters for PFG' and PFD'…*

$$\alpha_{PFG_1} LI'(t) - \left\{\alpha_{PFG_1} + \alpha_{PFD_2}\right\} IS'(t) - \left\{\alpha_{PFG_3} + \alpha_{PFD_2}\right\} PFS'(t) = 0$$

Preliminary Results

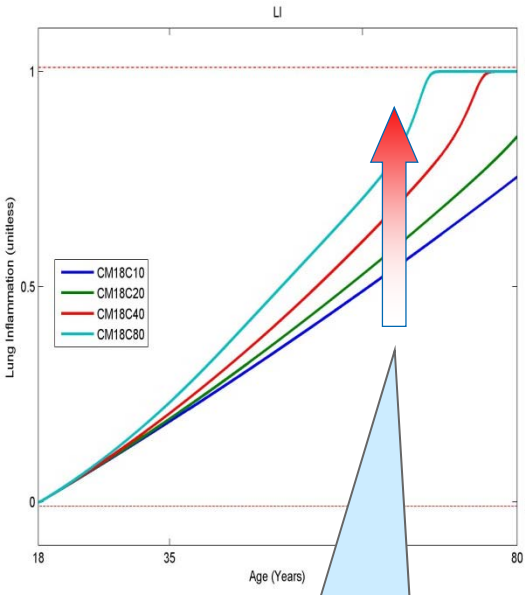
- Preliminary version of the model
  - Data acquisition is ongoing
  - Parameters were “guesstimates” based on biological experts’ opinions and expected behaviors of the system (model is not calibrated yet)

- Smoking scenarios used
  - Caucasian male, continuous smoking from ages 18 to 80 years old
  - Continuous smoking of 0.5, 1, 2, 4 packs/day

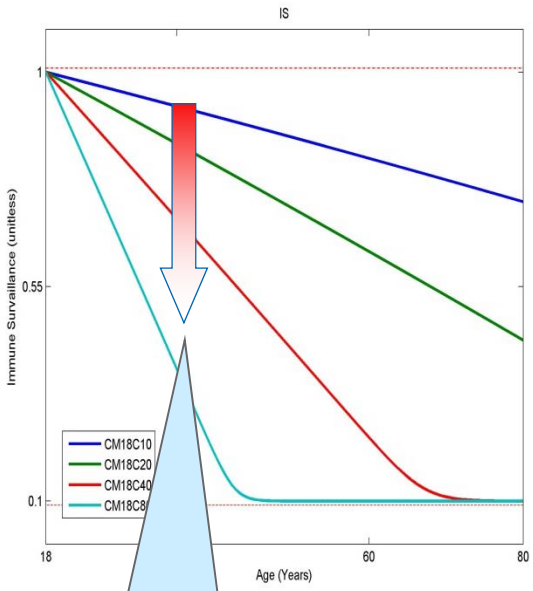
- ***Simulation results are illustrative and based on a non-calibrated model.***

Preliminary Results

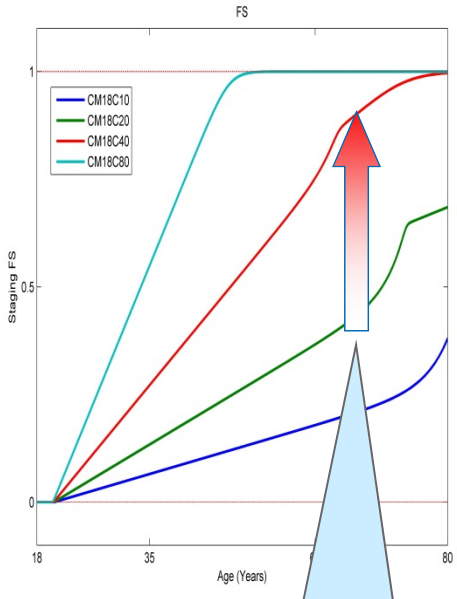
- Smoke dose accelerates the lung degradation…



Lung inflammation becomes more chronic faster with increasing dose and prolonged consumption



Immune surveillance capacity deteriorates faster with increasing dose and prolonged consumption



Field tissue progression (staging) is increasing with both dose and time

PMI RESEARCH & DEVELOPMENT

Summary

- Modeling approach was presented
  - Conceptual representation of the main biological steps (tradeoff to limit data requirements)
  - Translation of biology into mathematics
  - Data acquisition strategy and surrogacy
  - Optimization strategy to calibrate the model
- Model building and data acquisition are work in progress
- Current challenge: sufficient data of adequate quality…
  - Need data that focuses on initiation events, not on cancer or epidemiology (incidence or mortality)
  - Need data that has fully quantitative smoking-related information
  - *e.g. "Smoker", "Ex-smoker" instead of precise quantitative information on duration, dose, cessation period, problematic with pack-years data, …*

Thanks for your attention!

PMI RESEARCH & DEVELOPMENT