

# Lung cancer models and how to improve them

Olga Y Gorlova  
Baylor College of Medicine

October 11, 2019

# Types of models

## ◎ Prediction models

- Risk prediction models
- Early disease prediction models

## ◎ Natural history models

- Stage transition models
- Disease progression models

# Risk prediction (assessment) models

- Models predict risk of lung cancer within a specified time horizon
- Are based on:
  - Risk factors
  - Patient demographic characteristics - age, gender, ethnicity, education levels - at least some of them - are usually included

# Typically used risk factors

- Smoking history (at various levels of detail),
- Exposures: asbestos, dust, radiation;
- Physical activity
- Respiratory conditions, BMI, personal cancer history - can be assessed or self-reported);
- Family history of cancer (many varieties)
- Additional factors that are more difficult to assess: genetic variants individually or as a polygenic score, phenotypic assays of DNA repair capacity or mutagen sensitivity, fasting glucose levels, etc

# Examples of risk prediction models

- Bach model (original (2003) - first LC risk prediction model - and extended to include CT findings (2011)); 1-year risk that can be assessed recursively; developed from the data on ever smokers with high tobacco exposure (CARET); includes smoking characteristics and asbestos exposure; **AUC 0.66-0.72**
- Spitz model (original 2007 and extended to include phenotypic assays data 2008) based on a case-control study and SEER data; separate formulas for current, former, and never smokers, with different predictors; 1-year risk calculations can be applied recursively; includes smoking characteristic as appropriate; asbestos, dust, hay fever, emphysema, cancer family history; **AUC 0.57-0.69**
- Liverpool Lung Project (LLP) model (2008) - assesses 5-years risk and is used to enroll into the UKLS; includes smoking duration, personal history of cancer, pneumonia, lung cancer family history; **AUC 0.67-0.82**
- $PLCO_{M2012}$  model - most accurate of such models thus far; includes ethnicity, BMI, education, personal history of cancer, detailed smoking history, lung cancer family history, and COPD; assesses 6-years risk and is used to determine eligibility for LC screening in Canada; **AUC 0.803**

# Does adding genetic characteristics improve risk prediction?

- Adding select genetic variants improves the discrimination ability of models, but only slightly (Raji et al 2009; Li et al 2012)
- Adding markers of DNA repair capacity and mutagen sensitivity (believed to be genetic) likewise provides only a slight improvement (Spitz et al 2008)
- Using a comprehensive GWAS-derived polygenic risk score - adds little to discrimination (ILCCO work in progress)

# Applications of risk prediction models

- Currently the purpose is to better define the population in which lung cancer screening would be appropriate
- Trying to improve on the NLST eligibility criteria / USPSTF recommendations that are based solely on age and smoking history
- Less than 50% of lung cancer cases would have been eligible for screening by these criteria

# Early (curable) disease prediction models

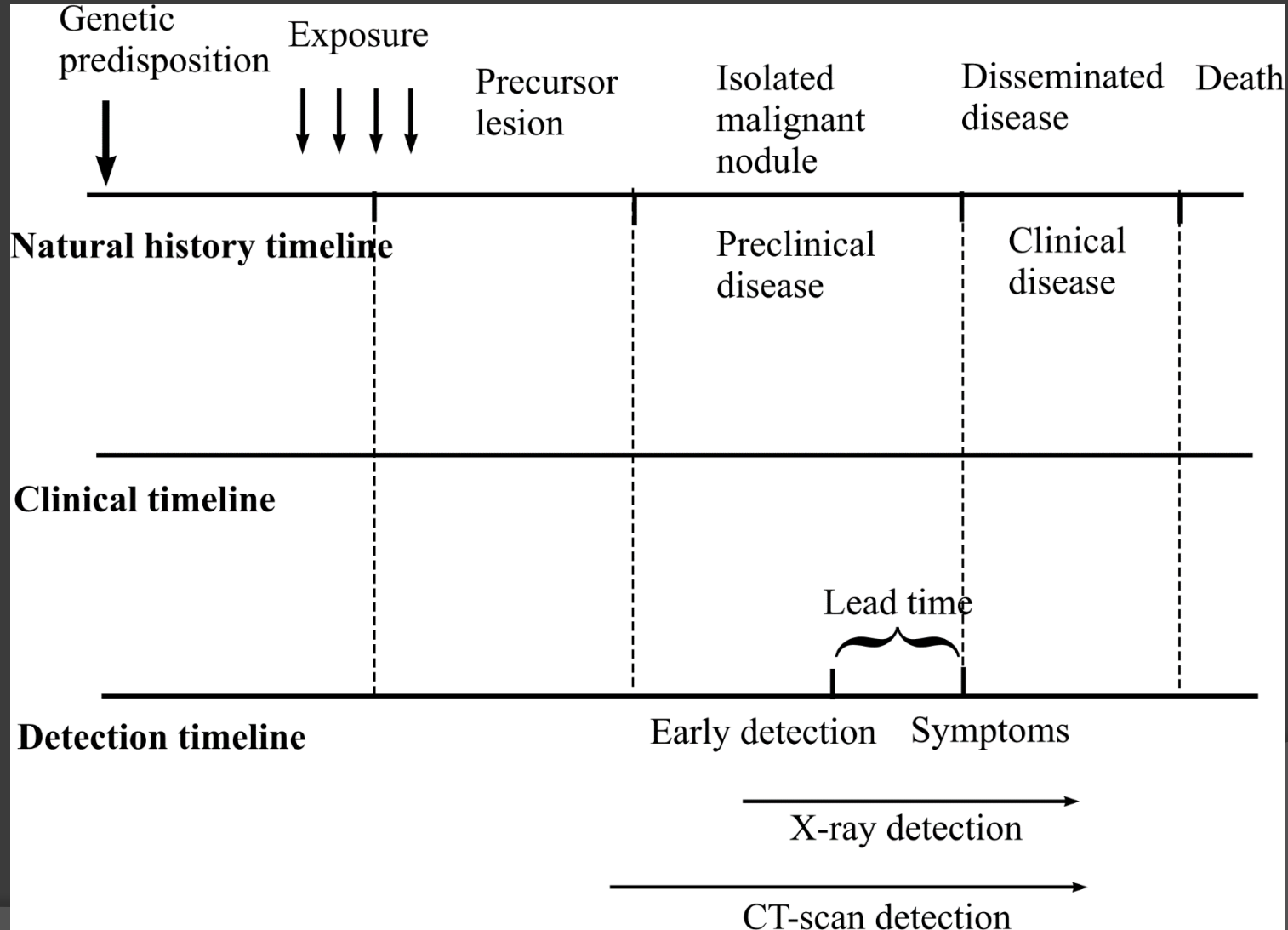
- None firmly established yet
- Proposed to be based on biomarkers of early disease (many different kinds: circulating proteins/miRNAs/circulating tumor DNA and mutation or methylation markers that can be detected in it/breath volatile organic compounds/radiologic or radiomics markers etc)
- The settings in which such models are to be applied are not very clear (e.g., before CT screen or after a screen with an indeterminate nodule?)



# Natural history models

- Reconstruct the natural history of disease
- Can represent state/stage transitions
- Or can model premalignant states followed by the formation of the first malignant cell, tumor growth, metastasizing, with associated symptoms, and death
- Disease detection can occur through symptoms or through periodic screening

# Lung cancer natural history



# Natural history models

- Deal with parameters which usually are not observable directly
- The values of these parameters are determined so that the model predictions are consistent with the observed data (usually from a trial)
- Then we can vary the parameters to explore hypothetical scenarios outside the settings of a fixed trial
- Notably it was modeling that first led to the conclusion about potential effectiveness of a highly sensitive modality for LC screening (Flehinger et al JNCI 1988)

# Lung cancer screening

- The National Lung Screening Trial (NLST) demonstrated lung cancer (LC) mortality reduction of 20% (over three rounds of screening).
- This led to a USPSTF screening recommendation.
- However, no generally accepted LC screening protocol currently exists.

# Optimizing a screening program

- The following may influence the performance of a screening program:
  - Eligibility criteria - define the population at risk
  - Screening interval - should depend on the typical progression rate of the disease
  - Definition and work-up of a positive finding - changes as screening experience accumulates
  - Duration of screening remains an open question.

# Why is continuous screening important?

- The absolute risk of lung cancer increases with age, even after accounting for competing risks.
- The lifetime lung cancer risk is substantial. E.g., for male current smokers it is ~25% if competing risks are eliminated but still ~15% if they are accounted for.
- Modeling can be useful in predicting the effect of continuous, long-term screening

# Modeling of continuous screening by the CISNET group (2014)

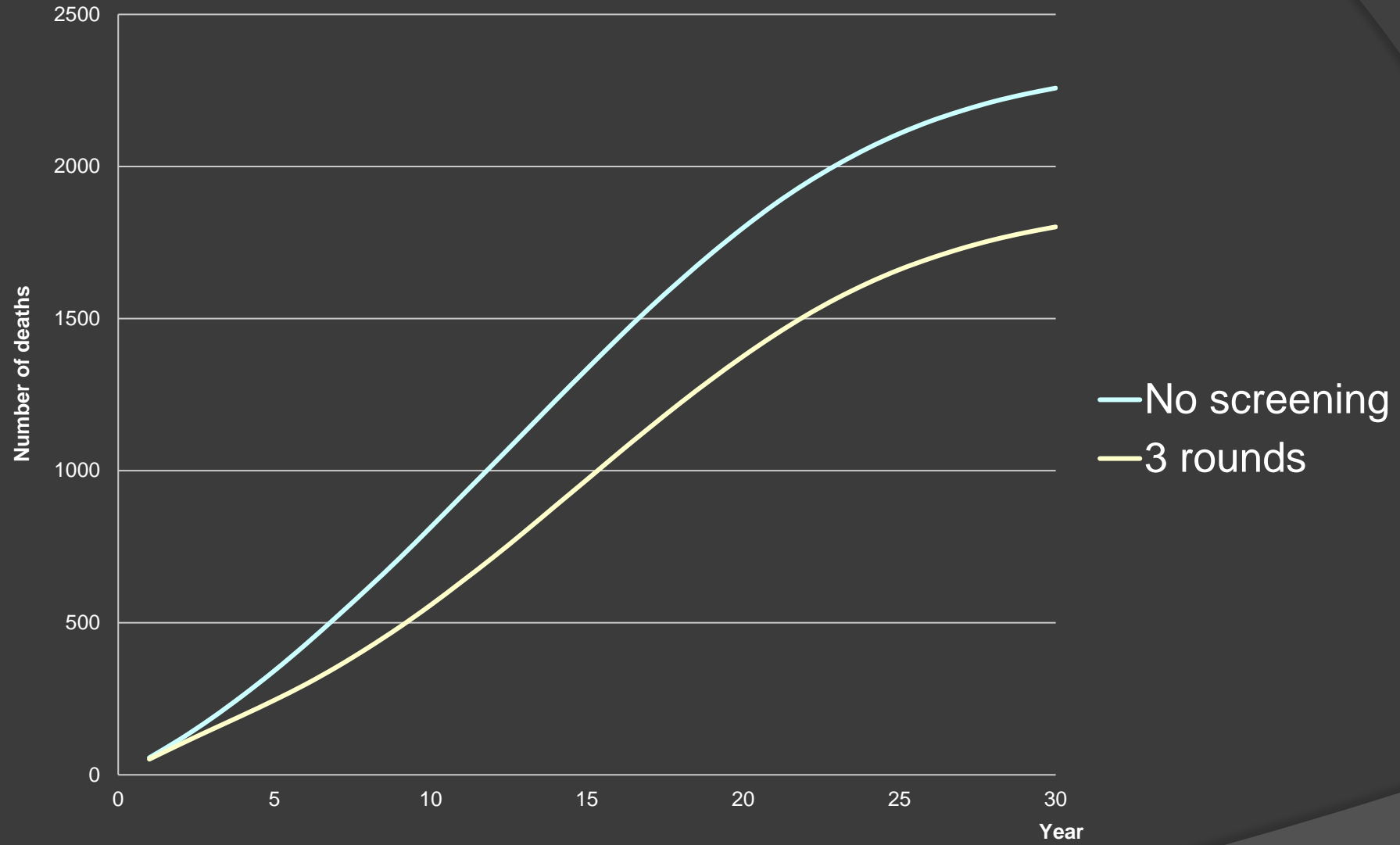
- Found annual screening from ages 55 through 80 most advantageous in terms of the number of lung cancer deaths averted
- Predicted 14% reduction in lung cancer mortality
- Models run without screening resulted in 37% of cases diagnosed with stage I disease (spontaneously), which may explain the low predicted mortality reduction

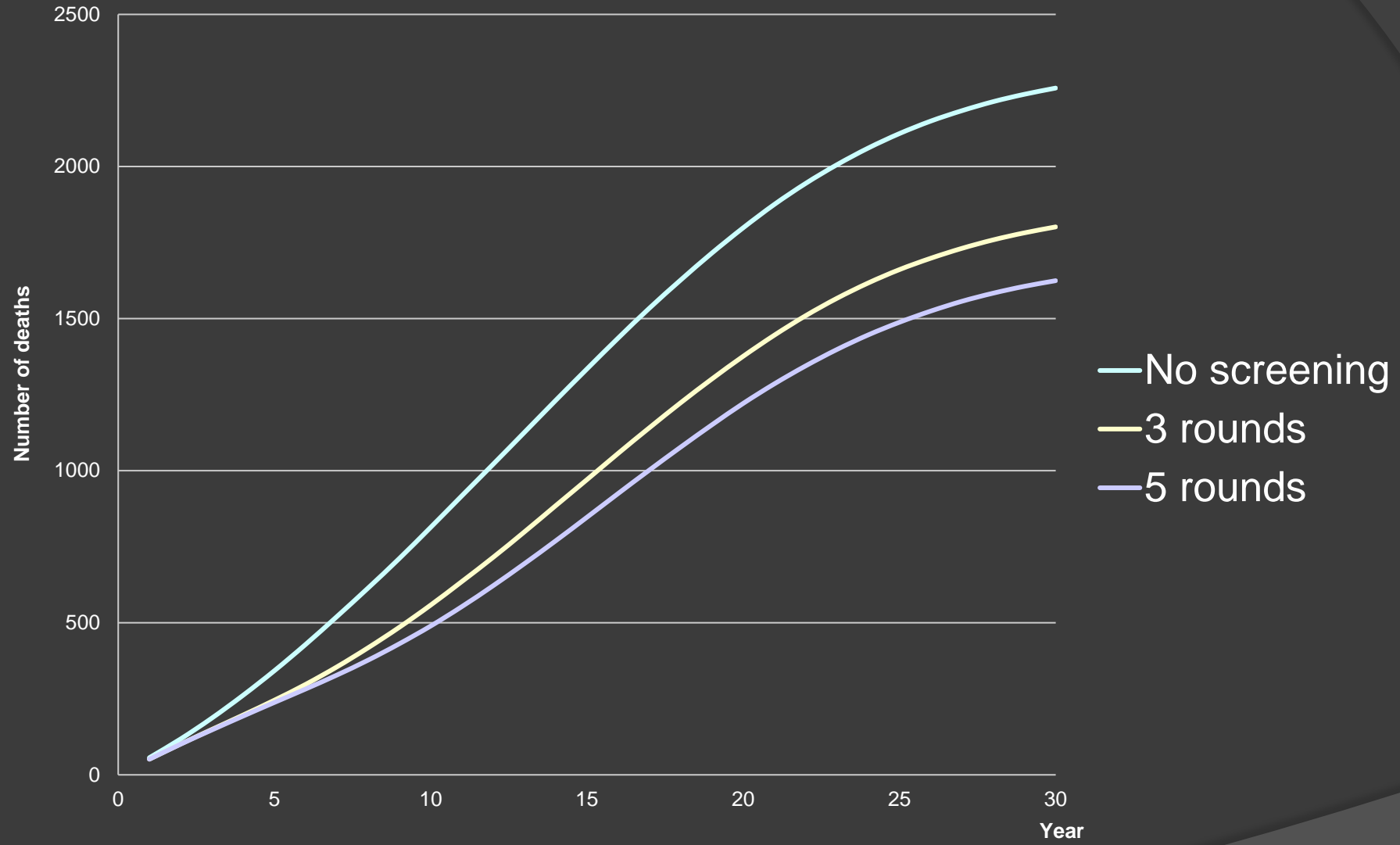
# The model

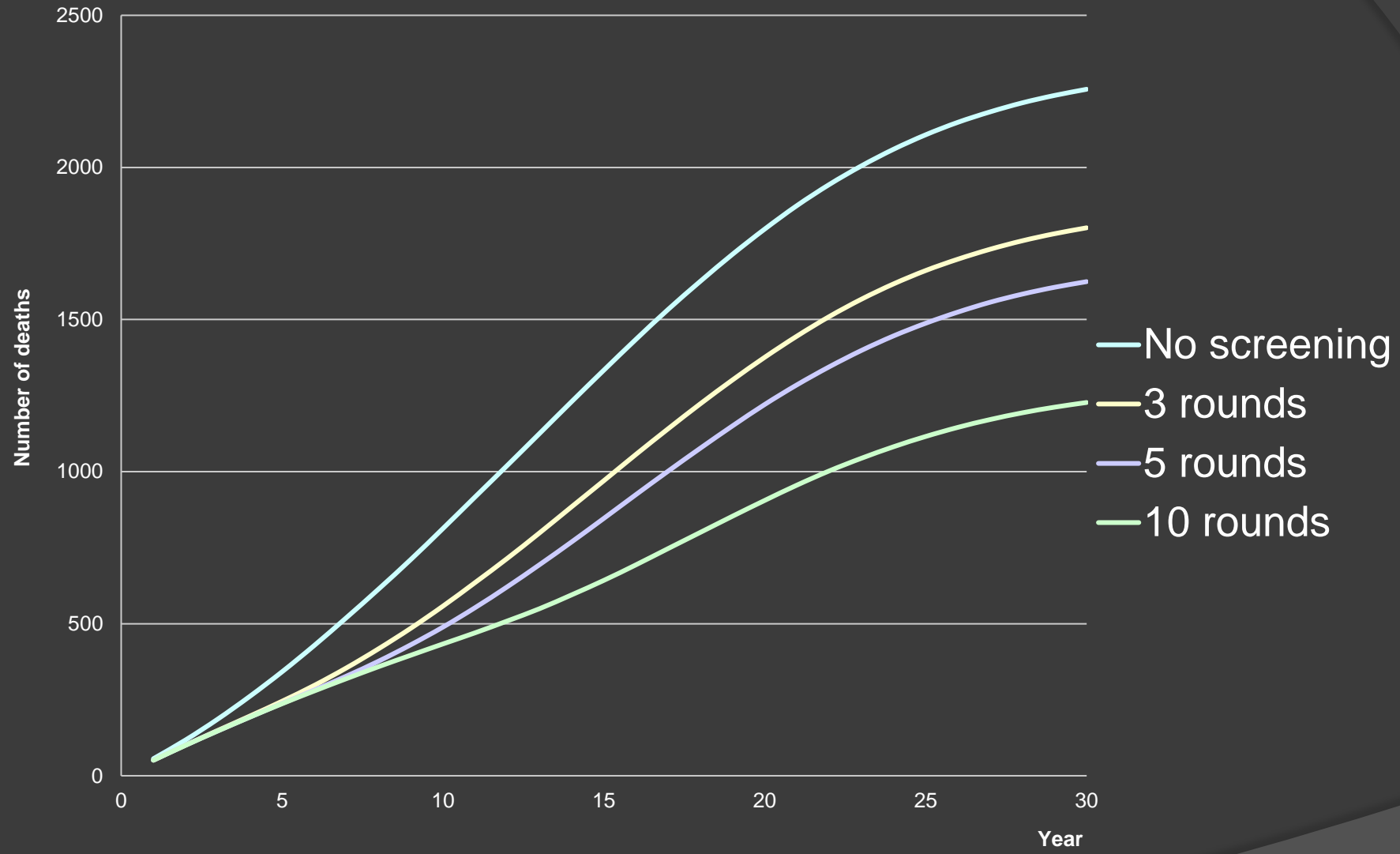
- We used our validated stage transition model of lung cancer natural history and screening (Gorlova et al 2001, 2003, 2005).
- The model was re-fitted to NLST data

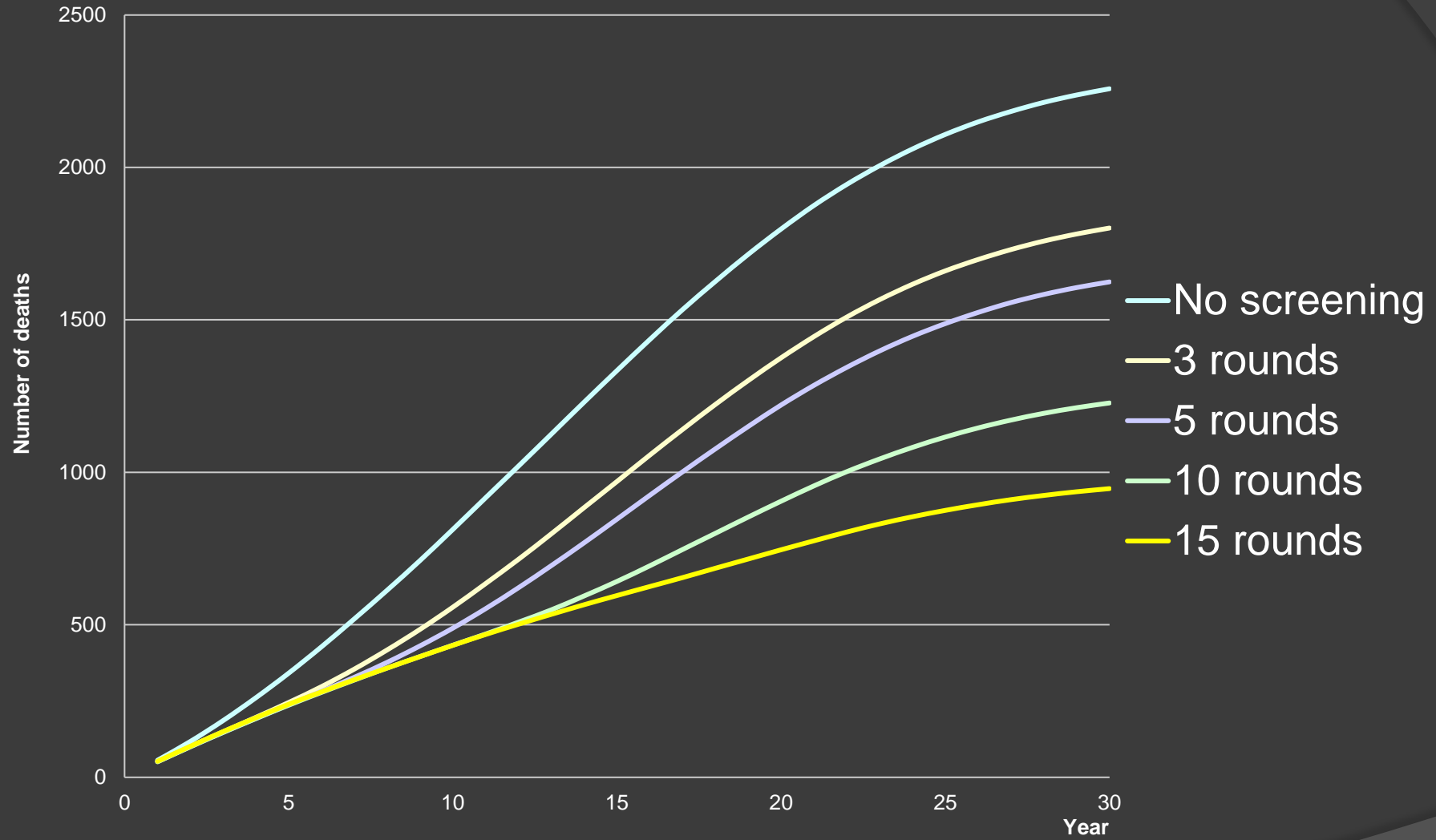


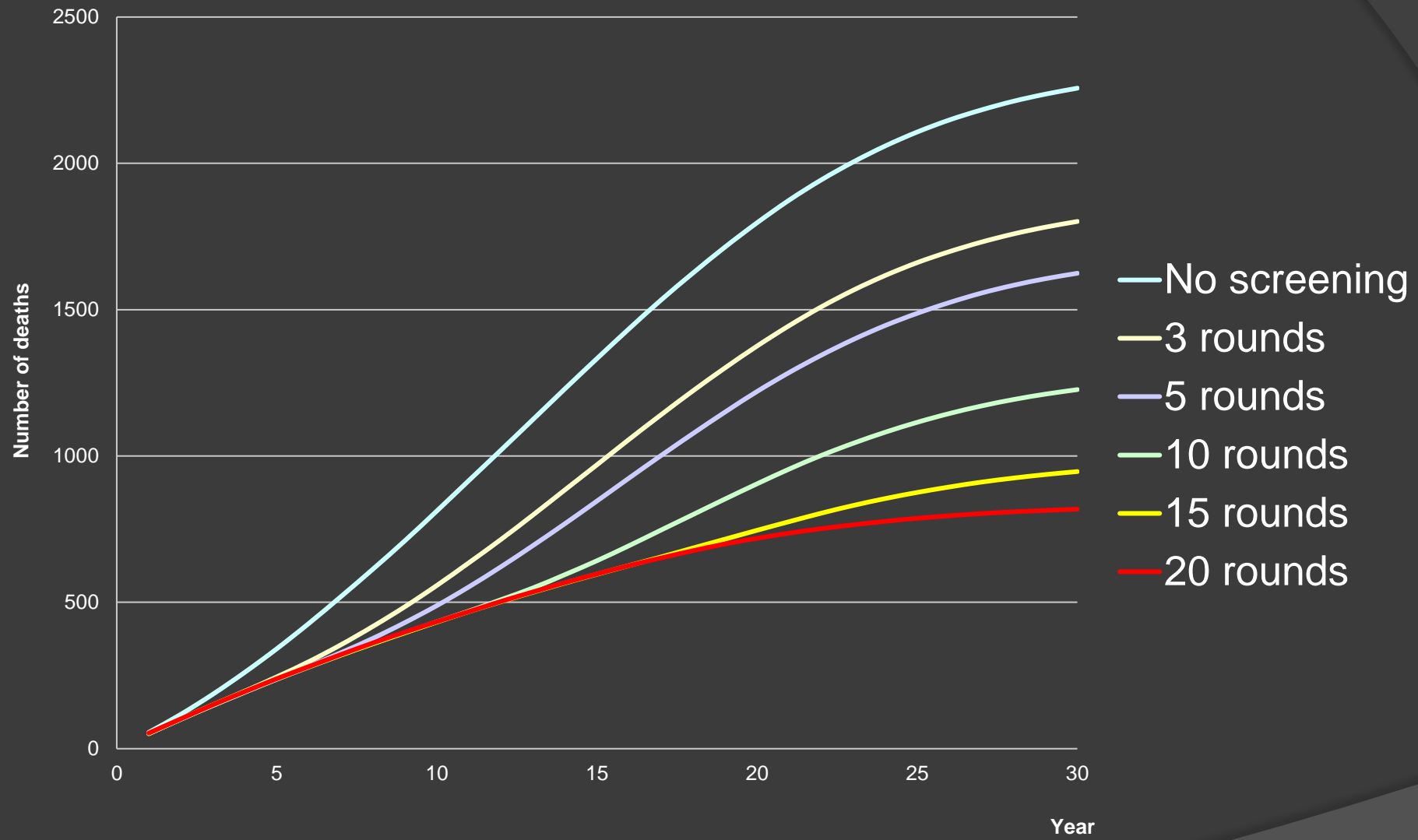
LC mortality under CT  
screening of different  
duration versus no screening

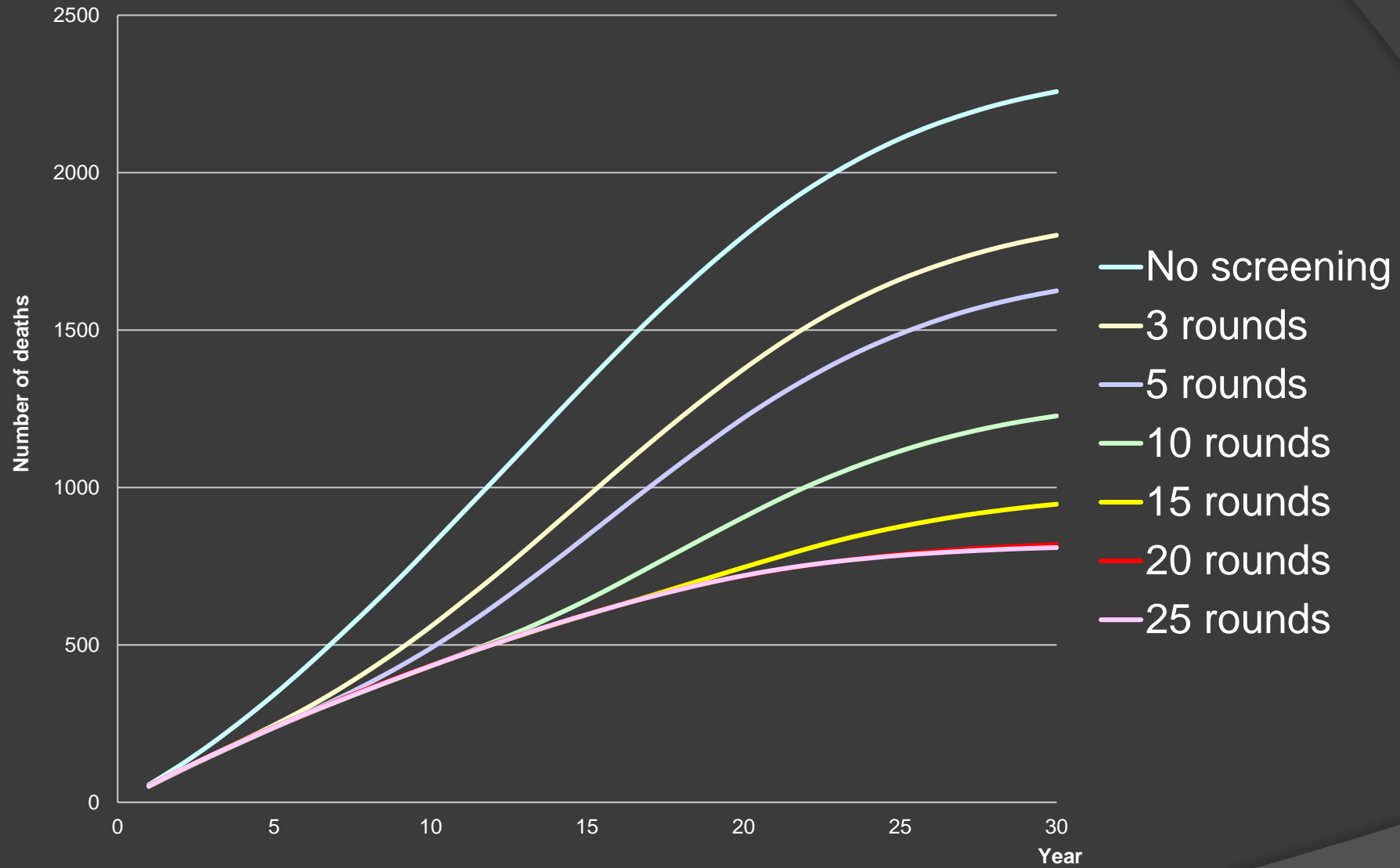


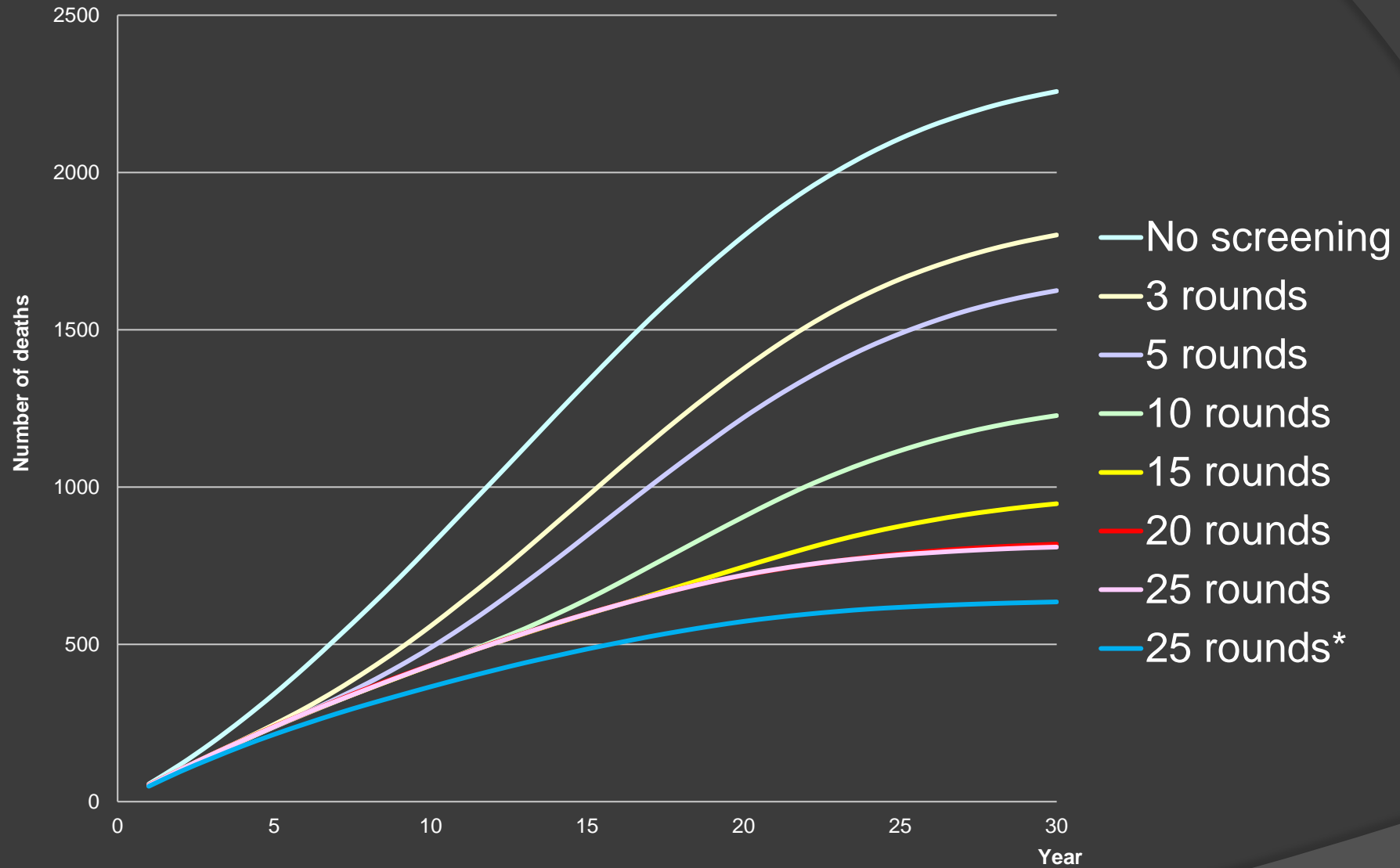






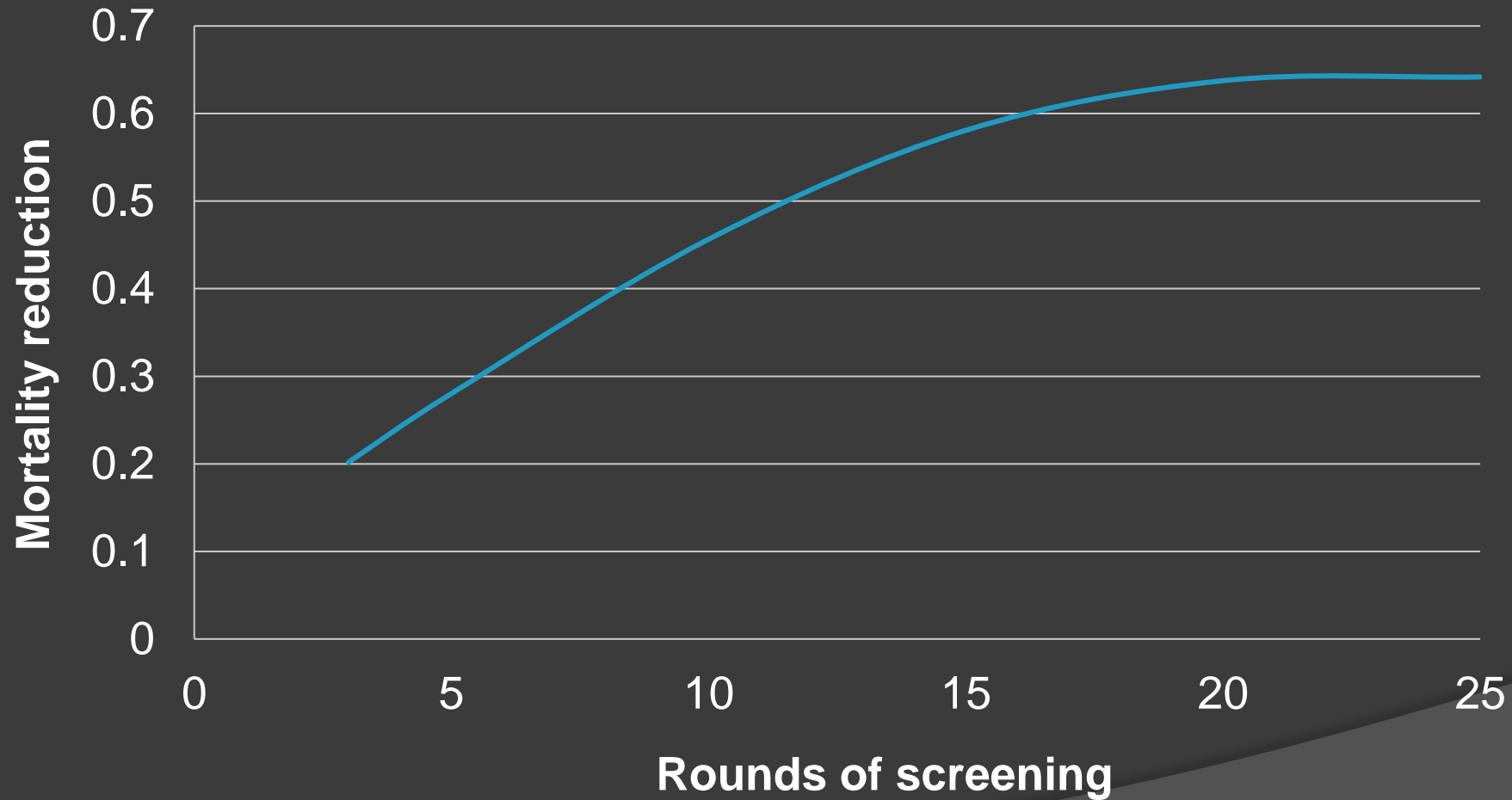








# Mortality reduction by the number of screens



# Conclusions from this modeling exercise

- The mortality benefit depends on the duration of screening
- Not surprisingly, the mortality reduction was higher with the continuous screening as opposed to a long-term follow-up
- Under the ideal circumstances, it can reach ~72%
- If screening starts at 55, continuing it for longer than 20 years carries little value

# Challenges (modeling and beyond)

# (Non-technical) challenges facing implementation and dissemination of lung cancer screening

- The lack of a clear understanding of potential benefits of LC screening by the physicians (Eberth et al Prev Med Rep 2018)
- The reduced enthusiasm and inadequate opportunity for participation in any health intervention program on the part of the highest risk group (heaviest smokers)
- The impact can be assessed by modeling

# Technical challenges

- Eligibility criteria (addressed by better risk prediction models)
  - Modeling of smoking cessation in the context of screening (Anderson et al CEBP 2009), with a corresponding lung cancer risk reduction
  - Modeling of a “healthy volunteer bias” to faithfully reproduce trial results
- Screening schedule: duration, frequency, possible changes in frequency
  - Natural history parameters (e.g. the typical growth rate distribution of CT-detectable tumors) should inform the schedule
  - Modeling of effects of risk factors on growth and progression rates may suggest personalized screening schedules

# Workup protocols

- Seldom considered by modeling
- Vary widely
- In terms of modeling call for:
- Malignant and benign nodules (size, number, consistency, dependence on risk factors)
- Incorporation of early disease biomarkers and an assessment of their value by modeling

# Missed opportunities

- Modeling of screening, with an appropriate subsequent intervention, not only for lung cancer, but jointly for conditions common in smokers (emphysema and coronary artery calcifications)
- Does the absence of emphysema call for less frequent LC screening?
- What magnitude of overall mortality reduction can be achieved?
- What about cost-effectiveness?

# Alternative directions: do symptoms hold any promise?

- Early stage lung cancer is believed to be asymptomatic and thus mostly found incidentally
- However, a chart review at MD Anderson showed that the top reason (52.5%) for the early LC patients to come to medical attention was symptoms (Chen et al 2012)



# Limitations of a chart review approach

- No controls
- Symptoms were reported at a single time point
- Can a more accurate recording of symptoms be useful?

# Possible alternative approaches

- Giving high-risk individuals a short questionnaire on a quarterly basis
- Creating a mobile app to log any unusual symptoms or changes in state, in the real time
- Developing a wearable device that can record and store/transmit the information on the respiratory function, in particular changes in cough (a prototype exists)

# Feasibility?

- The idea is that patterns of temporal changes in symptoms or objective measurements can serve for early detection
- Will need a lot of feasibility testing
- If successful, will generate a huge amount of data
- Will require new approaches to modeling

# Acknowledgements

Claudia Henschke  
David Yankelevitz

Marek Kimmel  
Millennia Foy  
Xing Chen