Innovative methods in Cancer Screening

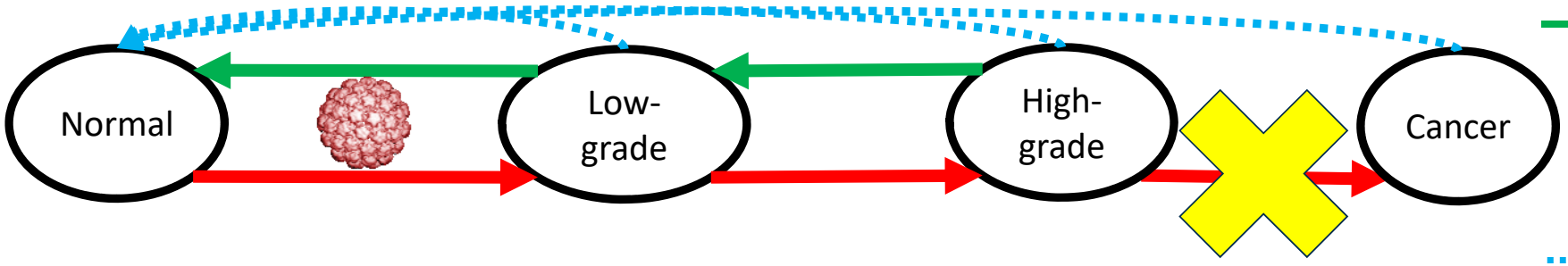
How Norway explores AI and uses the world's most powerful computer power to predict who has the largest risk of cervical cancer

Mari Nygård, MD, PhD
Head Research Department, Cancer Registry of Norway

Cancer Registry of Norway
INSTITUTE OF POPULATION-BASED CANCER RESEARCH

State

Transitions between stages

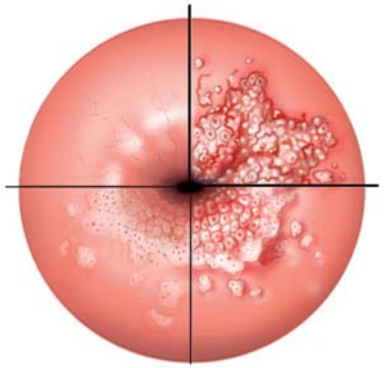


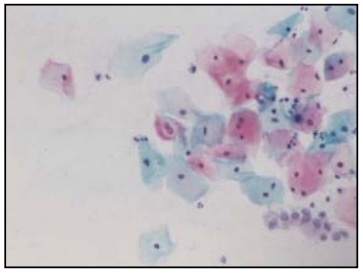
Progression

Regression

Treatment

Screening program

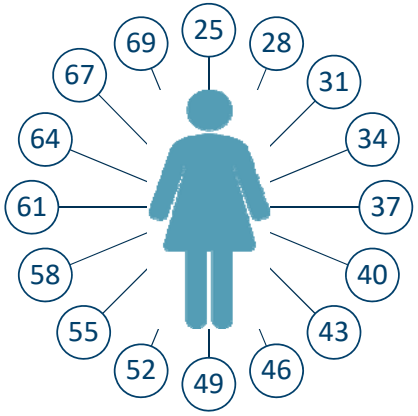




Mass-screening - a concept from 1960ies

Screening

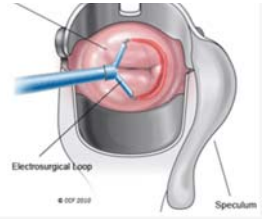
cytology exam for population without symptoms

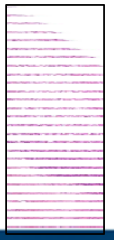


Diagnostic confirmation and treatment

Punch biopsy forceps

Loop electrosurgical excision procedure





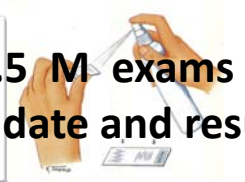
Mass-screening program - a concept from 1960ies was implemented in 1995 in Norway

ut symptoms

ID - women

1.8 M women in the database

date of the screening exam



About 12.5 M exams Including date and results

diagnose

ID- patient

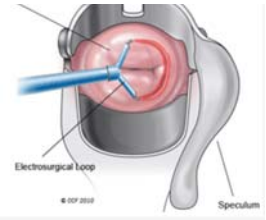
date of the diagnostic exam

Annually
about 10,000 Punch biopsies
About 6,000 treatments

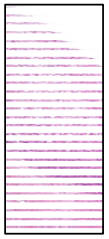
Punch biopsy forceps

date of the treatment

Loop electrosurgical excision procedure



diagnose



Mass-screening program - a concept from 1960ies was implemented in 1992 in Norway

ID - women

date of the screening exam

diagnose

Screening

cytology exam for population without symptoms

1.8 M women in the database

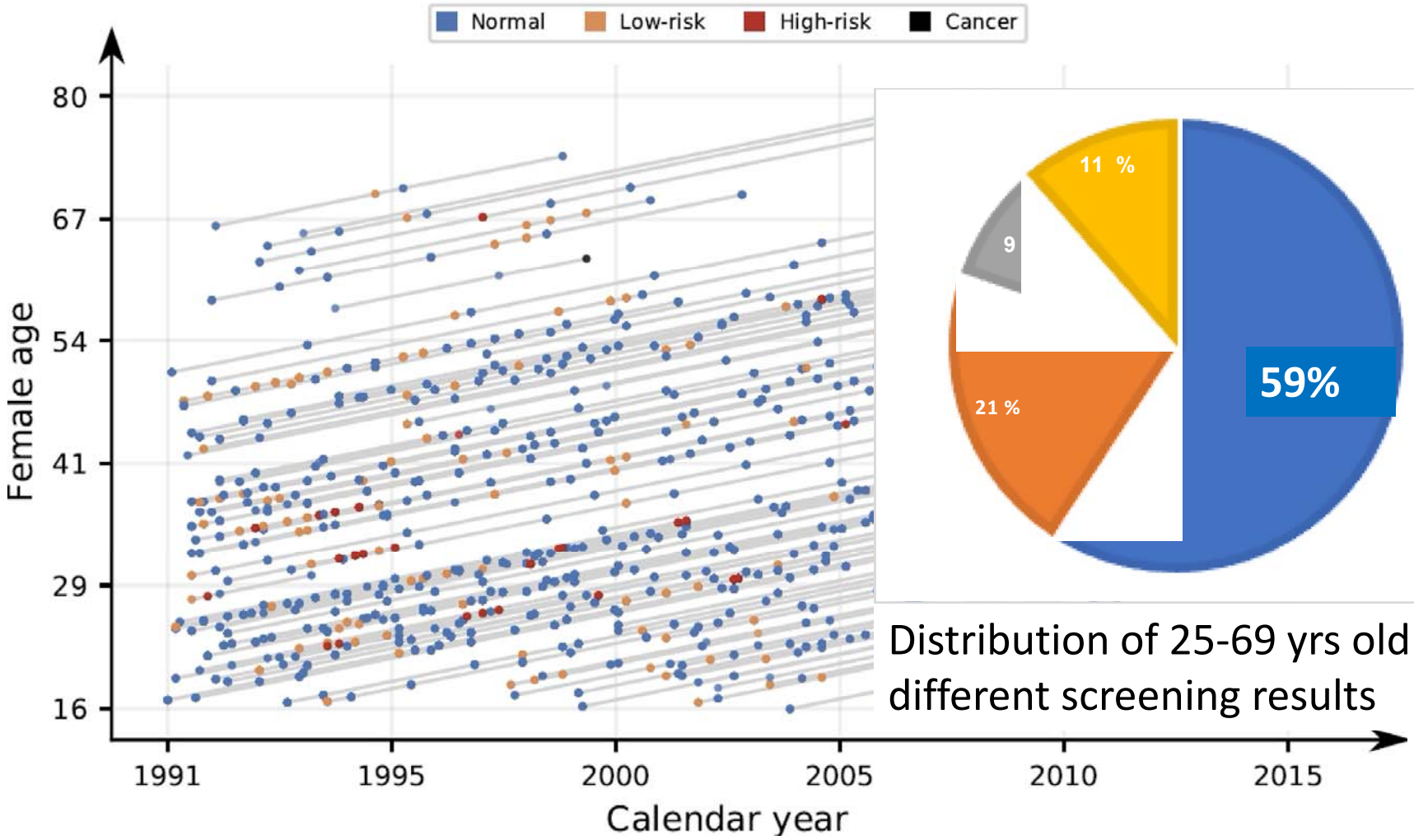
About 12.5 M exams Including date and results

90% of test are normal, 2% suggest disease, 5% suggest that “something is going on», 3 % are technical failures

Diagnostic confirmation and treatment

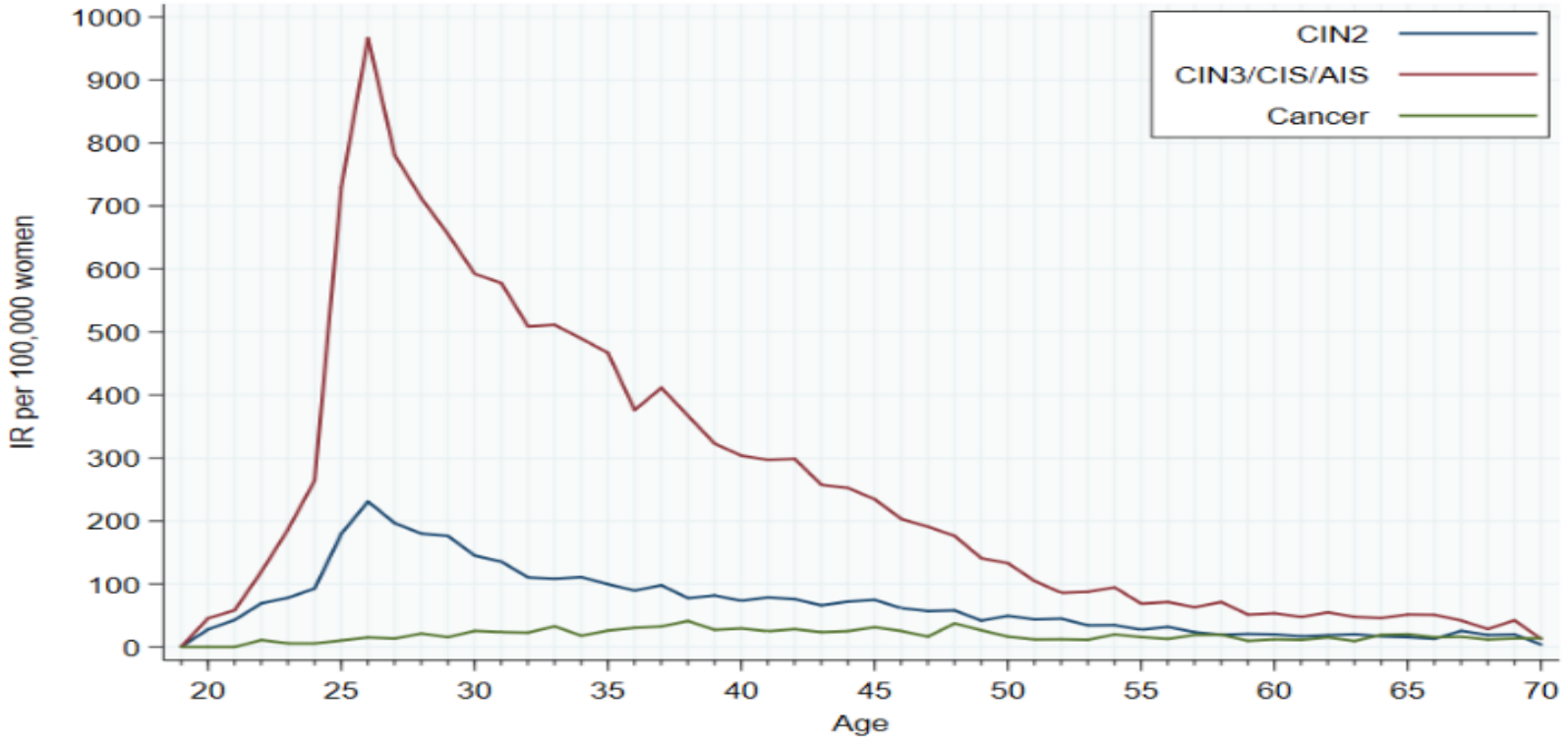
Annually
about 10,000 Punch biopsies
About 6,000 treatments

Including date and results



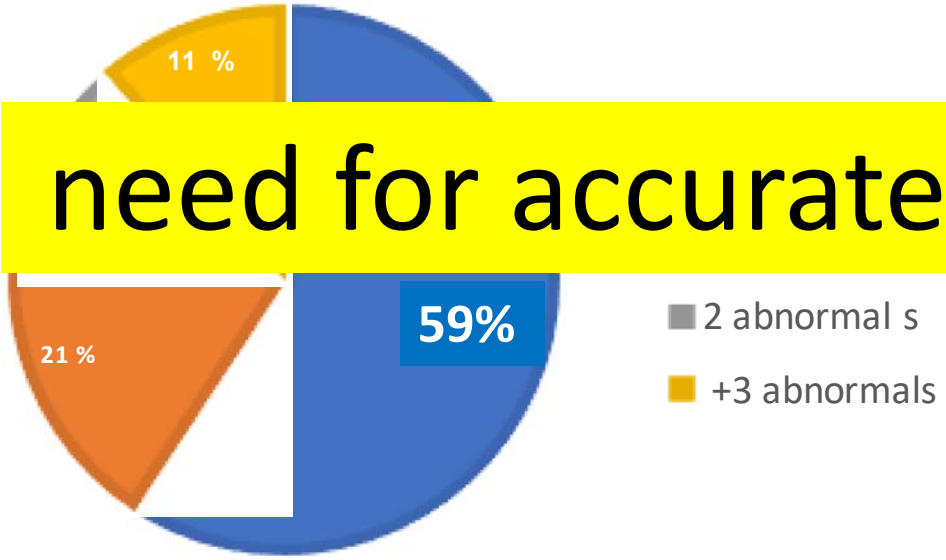
Distribution of 25-69 yrs old women with different screening results

Cervical pre-cancers and cancers by age in 2014-2016 Norway

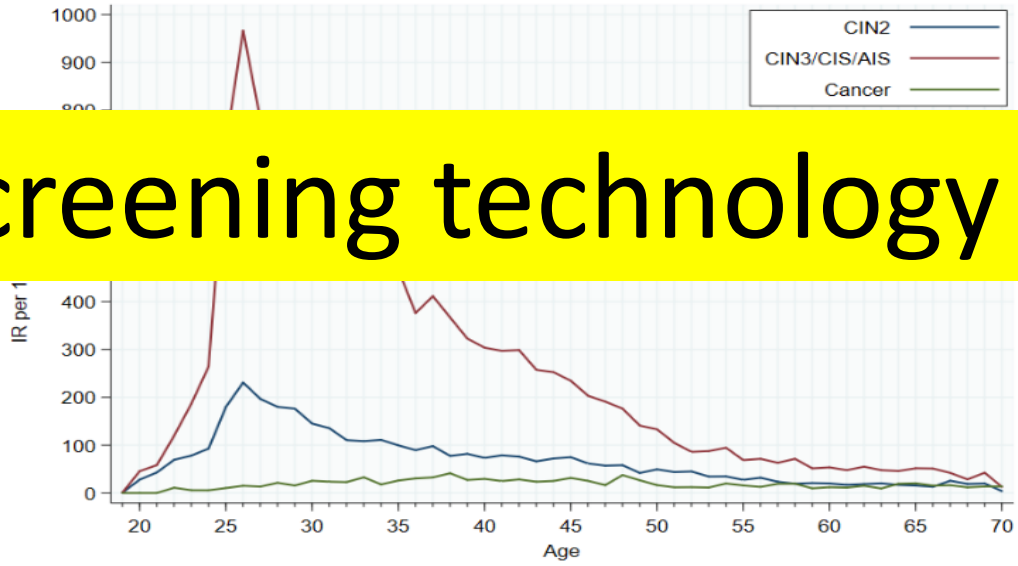


Harms of the mass-screening program: over-screening and over-treatment

need for accurate screening technology

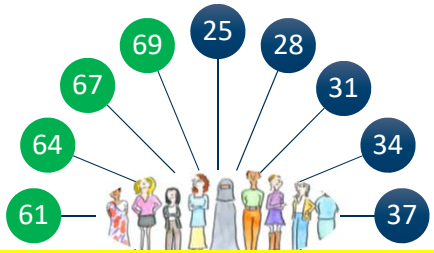


Distribution of 25-69 yrs old women with different screening results



Cervical pre-cancers and cancers by age in 2014-2016 Norway

Mass-screening program is not reaching to all women

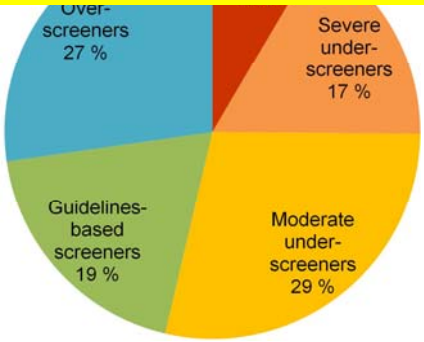


Proportion of those with one screening test in suggested time interval in Norway by age and year:

65% 25-39 yrs old

71% 40-54 yrs old

Screening should reach to those who need it most



Cancer incidence is highest among suboptimally screened

Pedersen et al. 2017 EJPH

Harmonizing Screening According To Risk

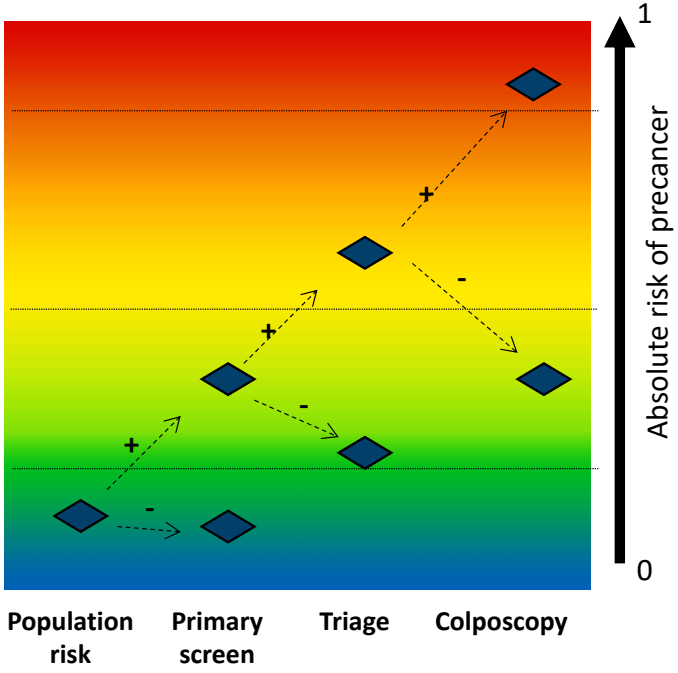
A

High risk:
Treatment

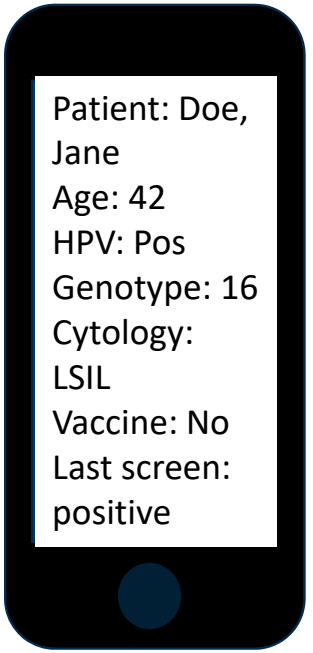
Medium risk:
Colposcopy

Low risk:
Triage or repeat testing

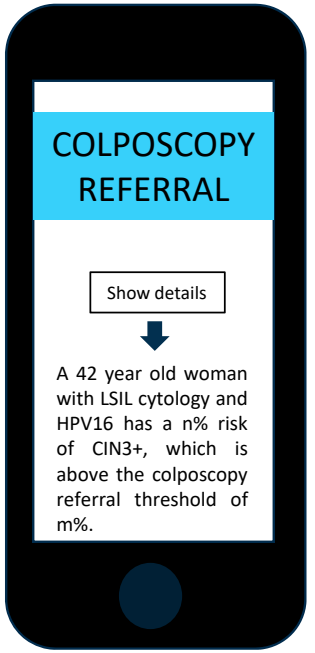
Minimal risk:
Regular screening interval



B



Data entry



Recommendation

Future Issues: Integration of Important Modifiers of Cervical Cancer Risk (personalized screening) Screening History, Integration of new biomarkers e.g., genotyping, p16 Ki67, mRNA, Methylation, HPV Vaccination

An HPC Application to Population-Level Cancer Screening Data

HPC and Data Intensive Resources at LLNL

Personal level data from the Cancer Registry of Norway

LLNL-PRES-753007

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Aim

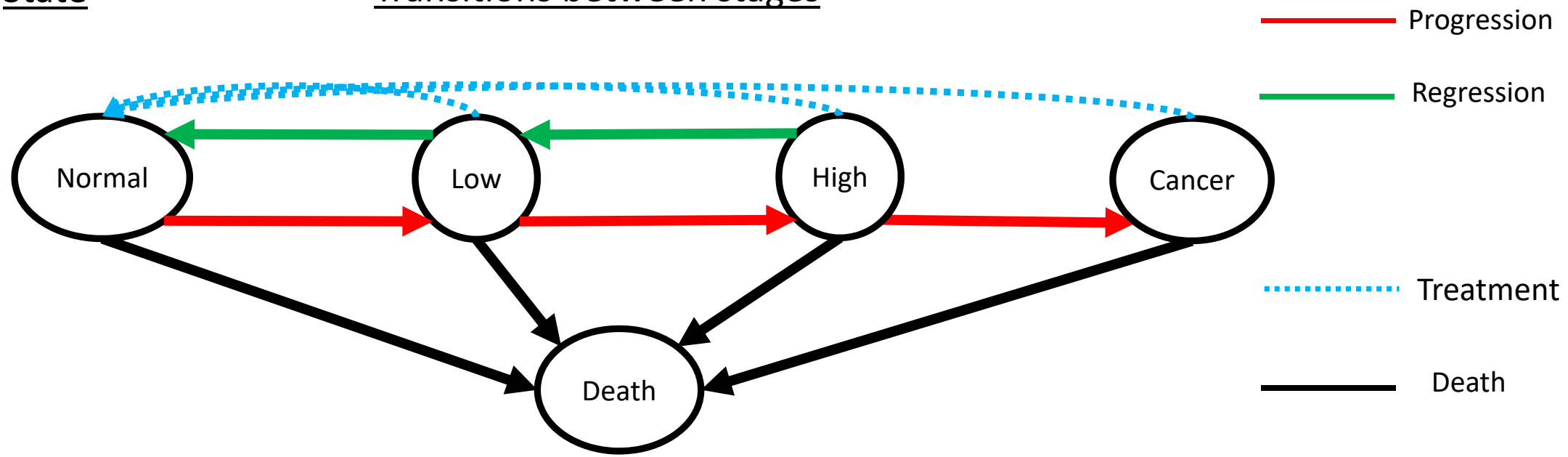
… to make cancer prevention programs more flexible, scalable and sustainable by use of the personalized algorithms.

If proven effective, our project will introduce a paradigm shift from mass-screening to a personalized prevention.

Hidden Markov Model - joint probability model

State

Transitions between stages



Model Screenings as Noisy Observations with Possible Misclassification

Soper et al. Statistics in Medicine, 2020

Continuous-Time Markov Process for Cervical Cancer

States

1 Low Grade
2 High Grade
3 Cancer
4 Death

Transitions

Progression
Regression
Death

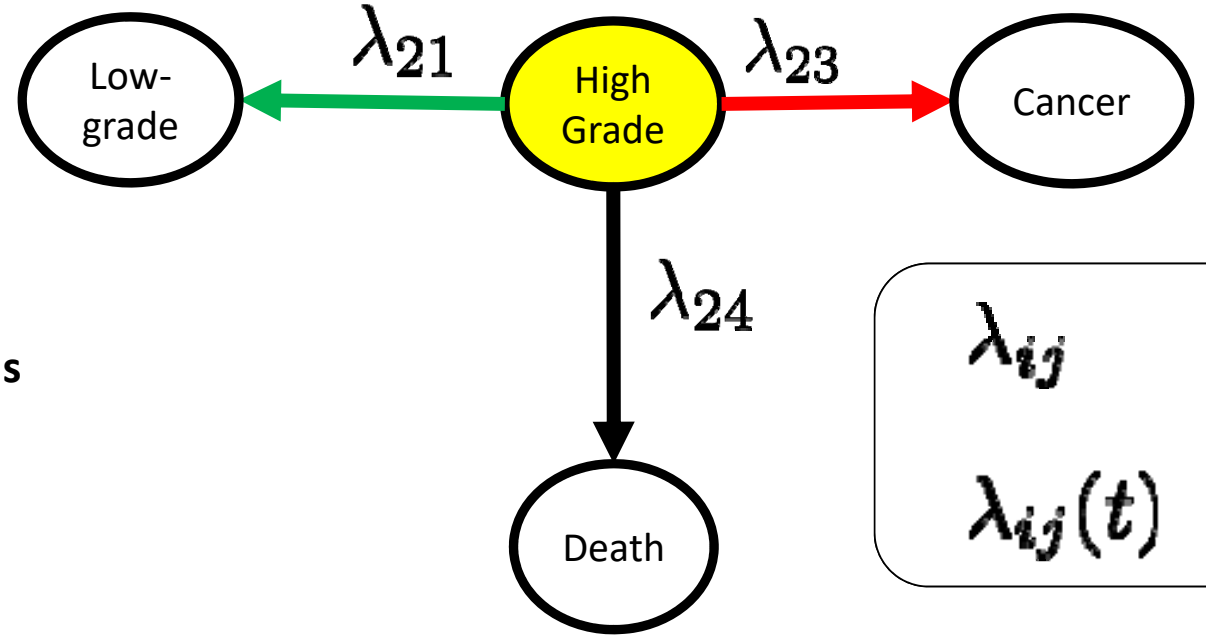
disease state always **unobserved**

S_{t_1}

Observed exams

X_{t_1}

Cytology
Histology
HPV



Transition intensities determine how likely it is to transfer out of a current stage and into another.

λ_{ij} ***Age-ignorant***

$\lambda_{ij}(t)$ ***Age-dependent***

Cancer Registry of Norway
INSTITUTE OF POPULATION-BASED CANCER RESEARCH

Soper et al. Statistics in Medicine, 2020

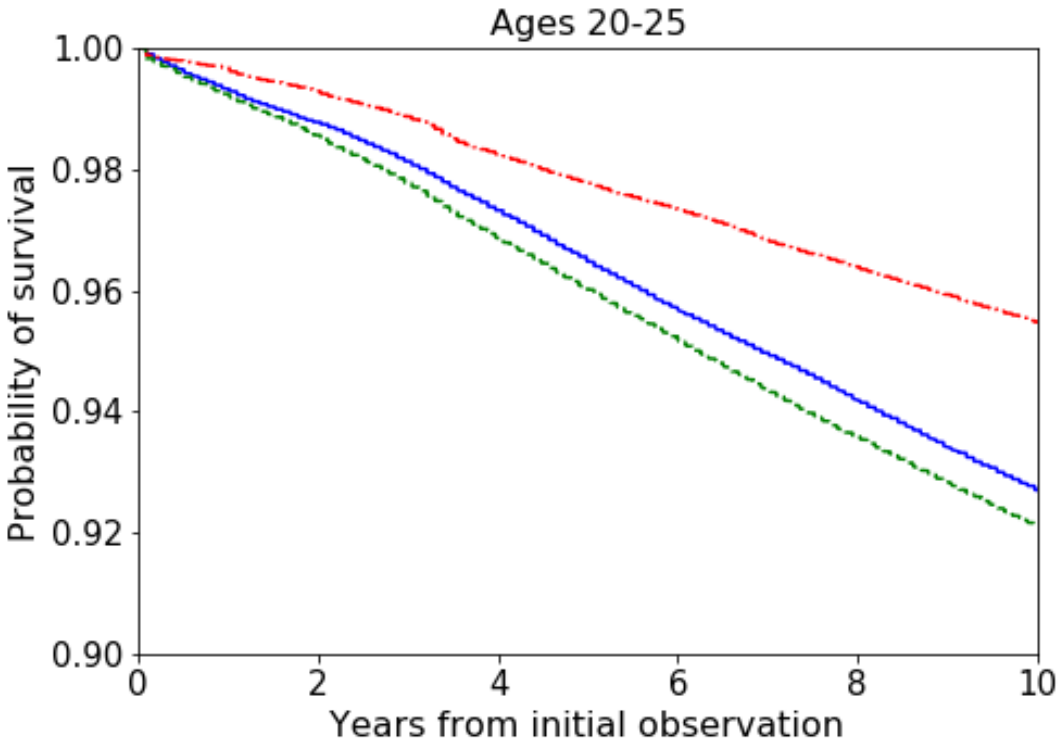
HPC and Data Intensive Resources at LLNL

- Catalyst, Quartz (Intel Ivy Bridge)
  - 7,776 cores across 324 nodes with 24 cores per node; 96,768 cores across 2,634 nodes with 36 cores per node
- Inference Algorithm
  - Implemented a modified EM algorithm in parallel to compute Maximum Likelihood Estimates (MLEs) of all model parameters.
- Simulations and Model Fit
  - Simulated synthetic population data from learned models

    Age-ignorant
    age-dependent model

  - Compute Kaplan-Meier estimators for different outcomes comparatively for observed and synthetic data to assess goodness-of-fit.

Simulated Data vs Observed Data transition from normal to high-risk



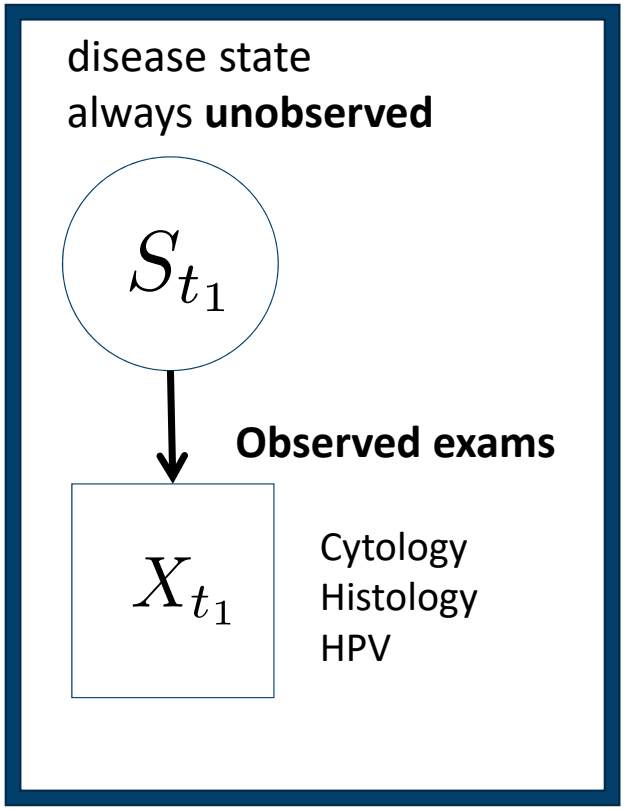
Soper et al. Statistics in Medicine, 2020

Learning likelihoods

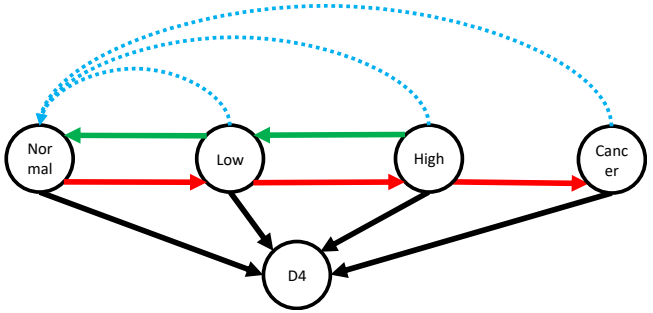
- Probability of cytology/HPV result conditioned on hidden state

	Test result: normal cytology	Test result: hrHPV positive
Normal state	99%	5.7%
Low-risk state	~0%	48.9%
High-risk state	12%	92.4%
Cancer	2.5%	98.8%

HPV test is more sensitive



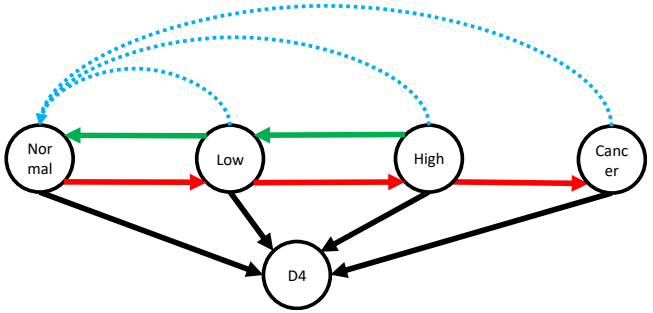
Learning likelihoods for transmission intensities



- Progression

Low-risk→High-risk 0.4% to 4.2%

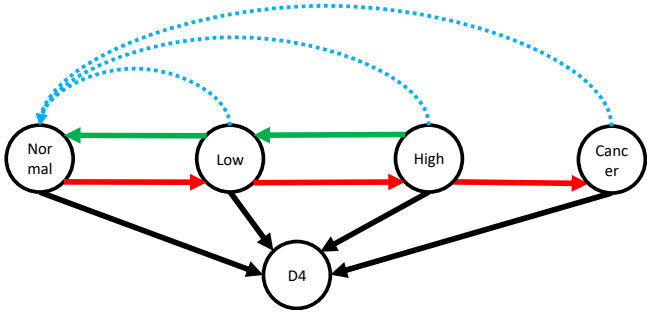
Learning likelihoods for transmission intensities



- Progression

Age group	20-24	35-39	60+
Low-risk→High-risk 0.4% to 4.2%	2.5%	3.5%	3.4%

Learning likelihoods for transmission intensities



Regression

Age group	20-24	35-39	60+
High-risk→Low-risk 5% to 22%	10.9%	6.7%	5.8%

WORK IN PROGRESS

Predict Future Disease States

- Compute probabilities for future disease states based on a patient's past test results.

$$P(S_t^{\text{new}} | \boldsymbol{X}_{t_1:t_k}^{\text{new}}),\ t > t_k$$

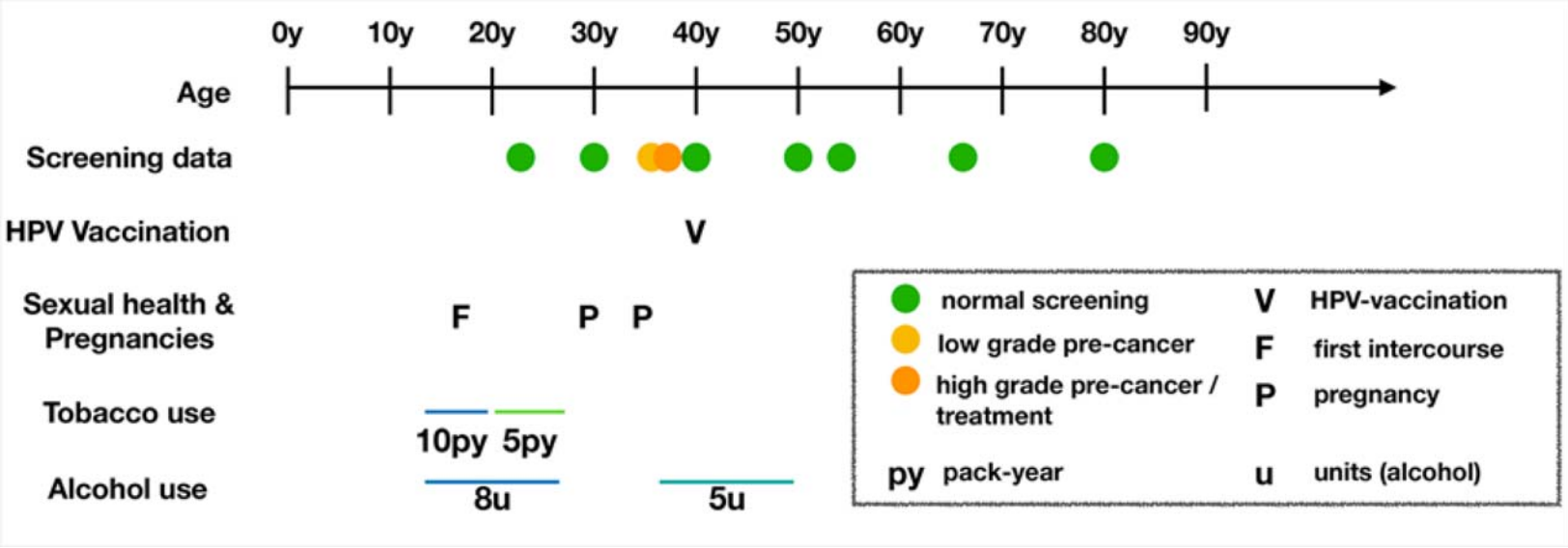
When should a patient be tested next?

Idea

Given a learned model, choose the minimum time interval that results in a positive diagnosis based on a Bayes optimal decision rule.

Increasing complexity: multi-modality datasets

- Detailed information on HPV types impacting the disease differently
- HPV vaccination status: age at vaccination, vaccine type
- Other possible life-style determinants (biomarkers)



2019 - expanding the group with PhD and Post-Doc

Cancer Registry of Norway
INSTITUTE OF POPULATION-BASED CANCER RESEARCH

Lawrence Livermore National Laboratory

The Research Council of Norway

Data-driven Framework for Personalised Cancer Screening

EEA Financial Mechanism 2014-2021 Baltic Research Programme

simulamet

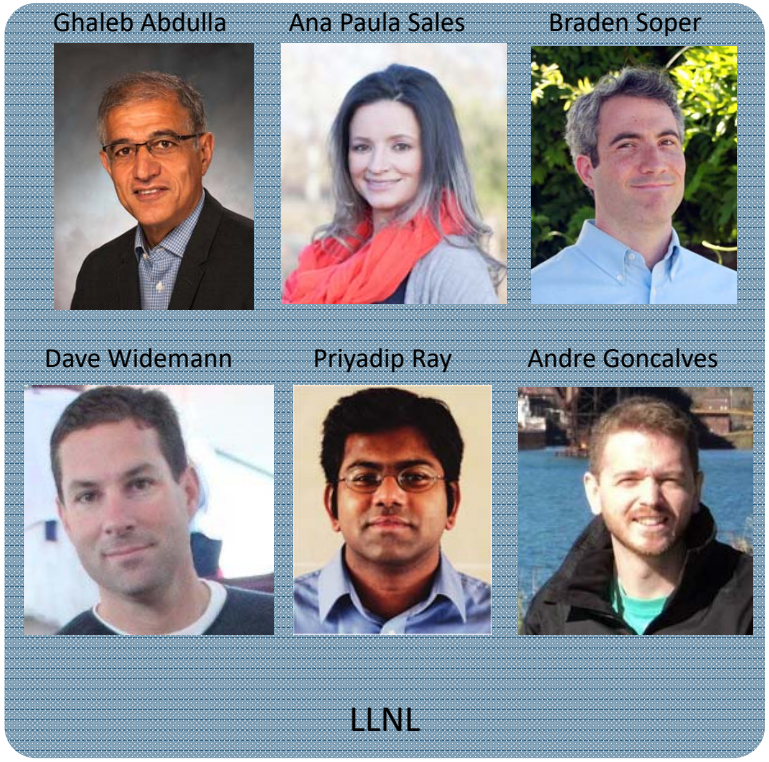
Karolinska Institutet

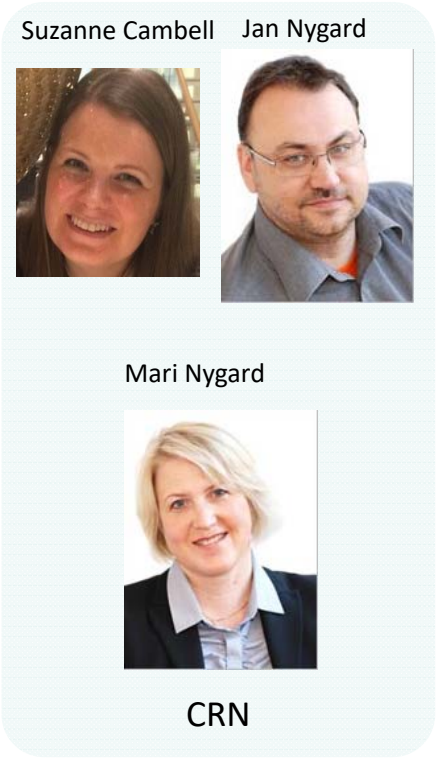
UNIVERSITY OF TARTU

Cancer Registry of Norway
INSTITUTE OF POPULATION-BASED CANCER RESEARCH

Lawrence Livermore National Laboratory

Our Team





- Supported
  - Giske Ursin (CRN)
  - Jim Brase (LLNL)
  - Rob Sharpe (LLNL)
  - Jason Paragas (LLNL)

Cancer Registry of Norway
INSTITUTE OF POPULATION-BASED CANCER RESEARCH

Collaborators

Braden Soper, LLNL

Ghaleb Abdulla, LLNL

Ana Paula Sales, LLNL,

Pryadip Ryi, LLNL

Mari Nygard, Cancer Registry of Norway

Jan Nygard, Cancer Registry of Norway

Hang Deng, Rutgers University

Cancer Registry of Norway
INSTITUTE OF POPULATION-BASED CANCER RESEARCH