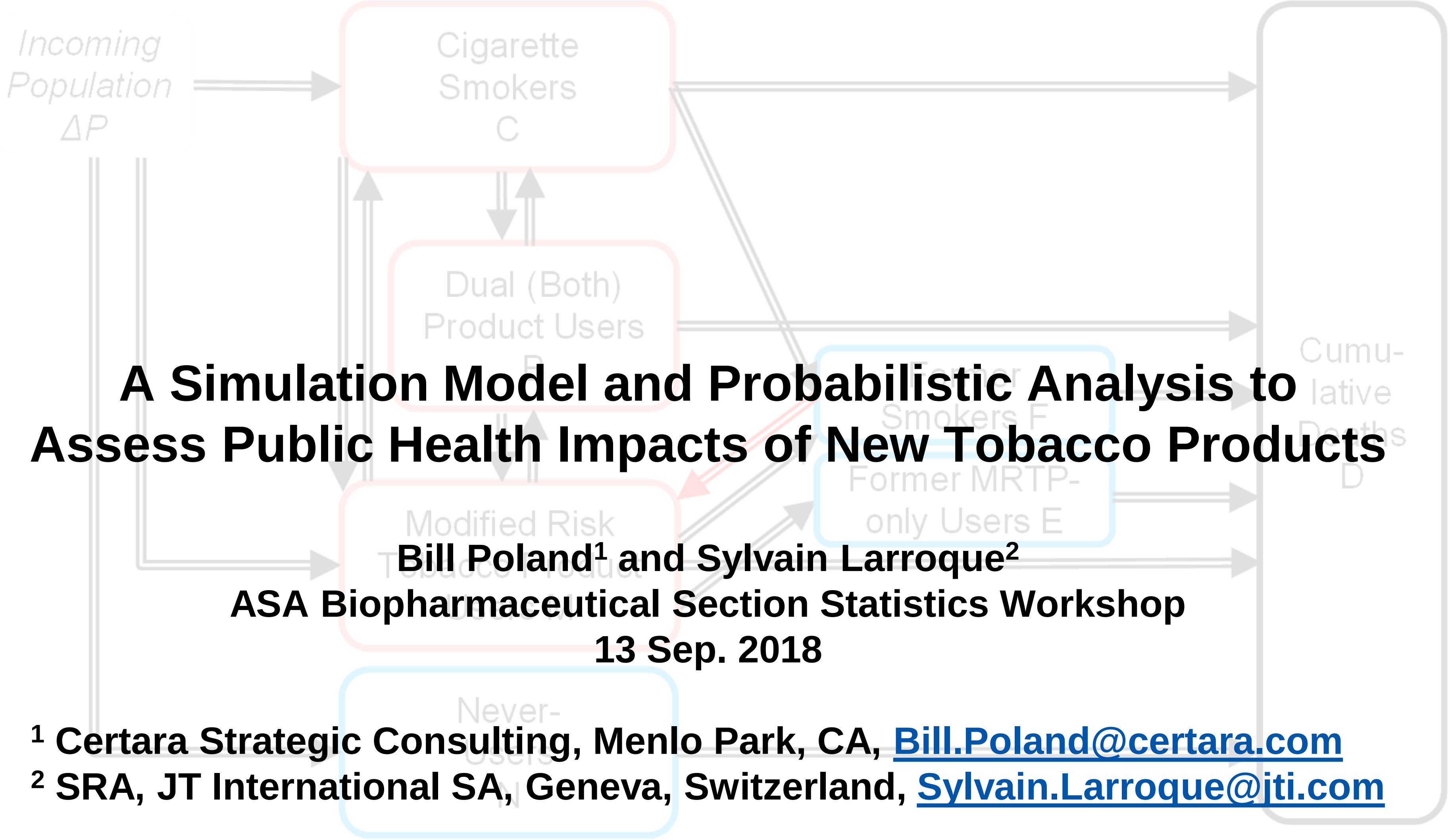


A Simulation Model and Probabilistic Analysis to Assess Public Health Impacts of New Tobacco Products

Bill Poland[1] and Sylvain Larroque[2]

ASA Biopharmaceutical Section Statistics Workshop

13 Sep. 2018

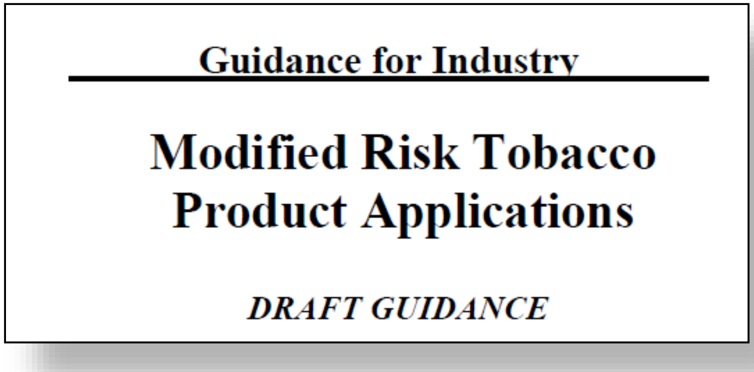
[1] **Certara Strategic Consulting, Menlo Park, CA, Bill.Poland@certara.com**
[2] **SRA, JT International SA, Geneva, Switzerland, Sylvain.Larroque@jti.com**

Agenda

- Population health modeling
- How the model works
- CPD-Excess Risk modeling
- Gradual Excess Risk changes
- Key outputs and sensitivity
- Probabilistic analysis

FDA encourages use of statistical models to project population health effects of a candidate Modified Risk Tobacco Product (MRTP).

- A Modified Risk Tobacco Product (MRTP) application submitted to the FDA must demonstrate that the MRTP will:
  - “A. **significantly** reduce harm and the risk of tobacco-related disease to **individual** tobacco users; and
  - B. benefit the health of the **population as a whole** taking into account both users of tobacco products and persons who do not currently use tobacco products.” http://www.fda.gov/TobaccoProducts/GuidanceComplianceRegulatoryInformation/ucm262077.htm
- “FDA encourages the development and application of **innovative analytical methods** to make preliminary estimates of the potential effects of some change in the marketplace.”
  2012 Draft Guidance for Modified Risk Tobacco Products, www.fda.gov/TobaccoProducts/Labeling/RulesRegulationsGuidance/ucm297750.htm



Guidance for Industry

Modified Risk Tobacco Product Applications

DRAFT GUIDANCE

- Models have been used to project impacts on public health of new tobacco products with varying risk levels, as well as policy changes.

Models for predicting population health impacts of tobacco products can be categorized by population and by method.

Population:

- **Cohort** models follow a single cohort such as a birth-year cohort through death or an old age.
  - Calendar time might or might not be used.
- **Full (cross-sectional over time) population** models follow new incoming population as well as a specified initial population (of all ages) over time.
  - Births, sometimes net immigration

Method:

- **System Dynamics** calculates counts over time of every population subgroup (e.g., by year of age, sex, tobacco use status, and calendar year); transition probabilities are not influenced by prior history (“Markov assumption”).
- **Microsimulation (individual Monte Carlo)** methods draw random characteristics of many individuals (“agents”) and random transitions from population distributions, and then calculate statistics of the simulated population.

Teischinger & Poland 2017, TSRC, 28 Nov.-1 Dec. 2017, Abstract #44. https://www.coresta.org/sites/default/files/abstracts/2017_TSRC44_Teischinger.pdf

Each category has strengths and limitations.

Population, Method	*Strengths*	*Limitations*
Cohort	Requires fewer inputs & assumptions	Harder to interpret and apply to actual & projected populations
Full	Projects an actual evolving population Easily calibrated against historical data Can reduce to cohort model by zeroing incoming population	Requires birth/incoming population assumptions Population at any given time could be biased relative to subgroups of interest
System Dynamics	Deterministic calculation: no random variability in results	Must limit subgroup count to keep model manageable
Microsimulation	Intuitive Easy to incorporate complex effects (e.g., Cig./Day effects)	Need large sample to smooth out variability in results

Teischinger & Poland 2017, TSRC, 28 Nov.-1 Dec. 2017, Abstract #44. https://www.coresta.org/sites/default/files/abstracts/2017_TSRC44_Teischinger.pdf

Our model also has particular strengths and limitations.

Strengths

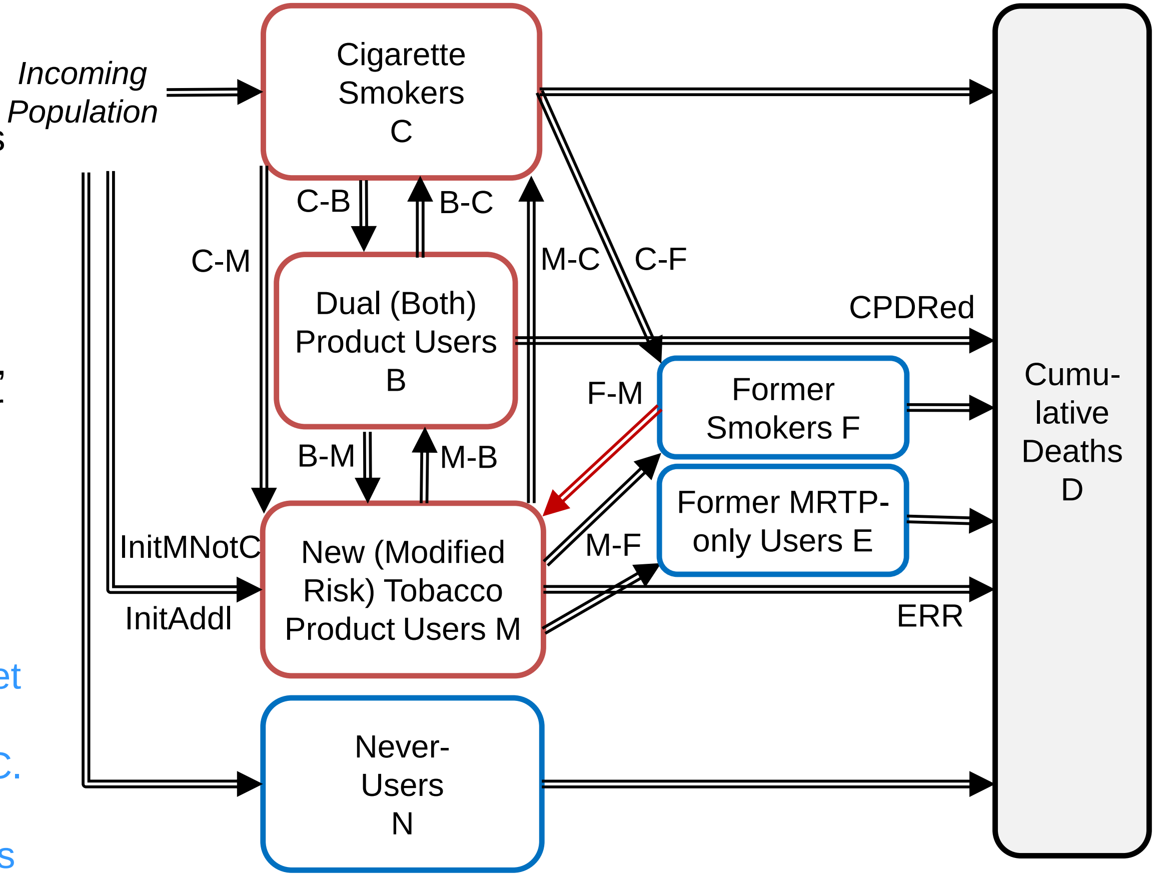
- Fast and easy to use
  - especially with good Excel skills
- Represents a full population evolving over time, and can analyze cohorts
  - by turning off incoming pop.
- Includes dual use, with CPD effects on mortality
- Allows discrete or continuous probabilistic analysis
- Easy to extend

Limitations

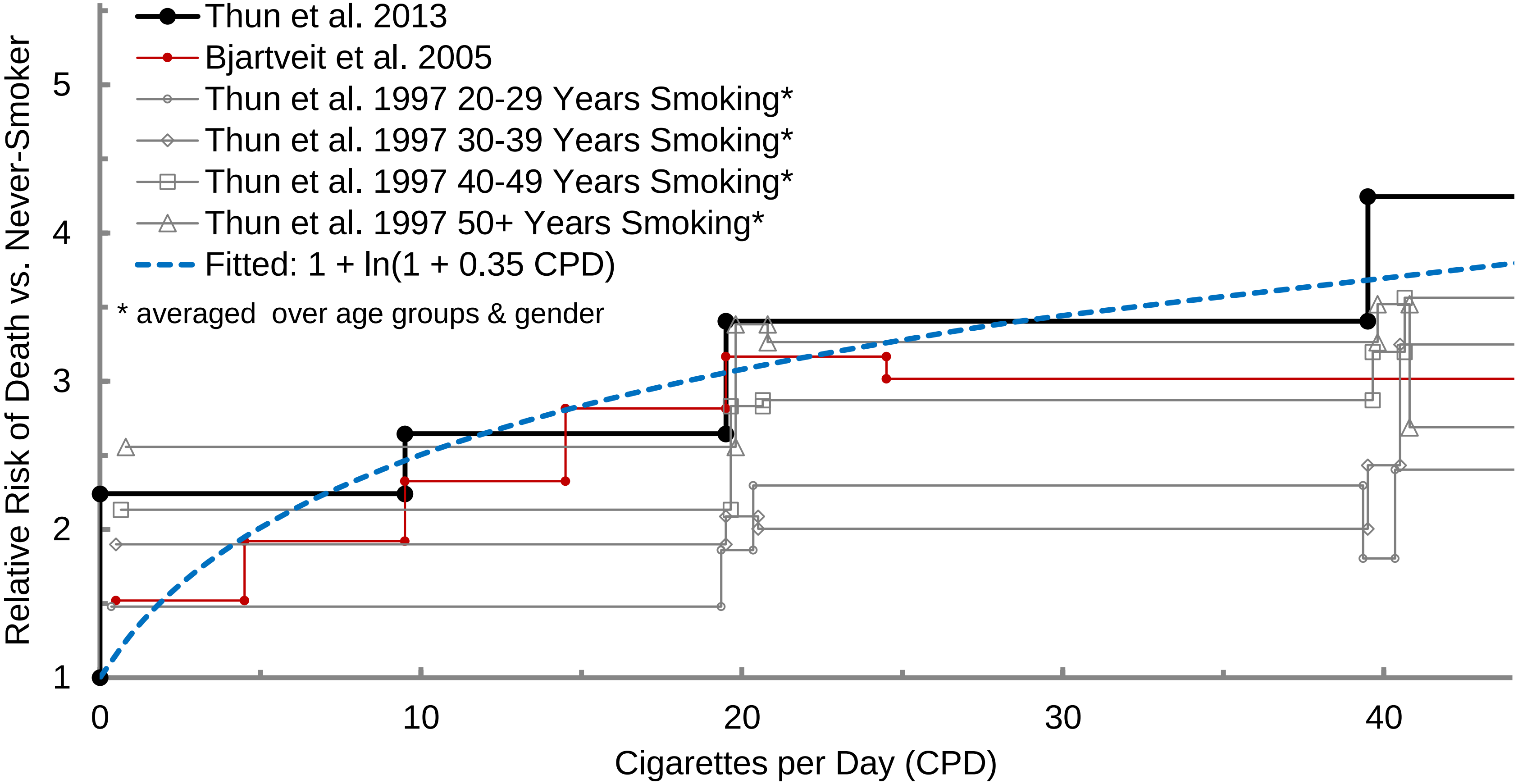
- Generic disadvantages of approach: requires demographic projections, attention to sample size
- Assumptions made for simplicity:
  - two products only
  - initiation at a single age
  - constant annual transition rates
  - quit only 1 product at a time
  - no relapse (permanent quitting)
  - use of mortality without morbidity or cost measures.

The simulation algorithm captures each transition in the model diagram.

- Loop to simulate each person *independently*.
  - In nested inner loop, step through each year of the simulation period.
- **Initialize** each person by random draws from input distributions for age, sex, initial tobacco use status if already ≥18 (C/F/N), and CPD for ever-smokers.
- Each year after the first, increment age and **check if the person dies** that year, based on a random draw accounting for age, sex, and tobacco use history.
  - If so, end simulation of that person.
  - Otherwise, update annual status based on random draws for all possibilities allowed for current state:
    - If just reached initiation age (18), set initial status to C, M, or N.
    - C becomes B, M, or F, or remains C.
    - B becomes C or M, or remains B.
    - M becomes C, B, or F/E, or remains M.
    - F becomes M, or remains F.
    - E and N remain the same.



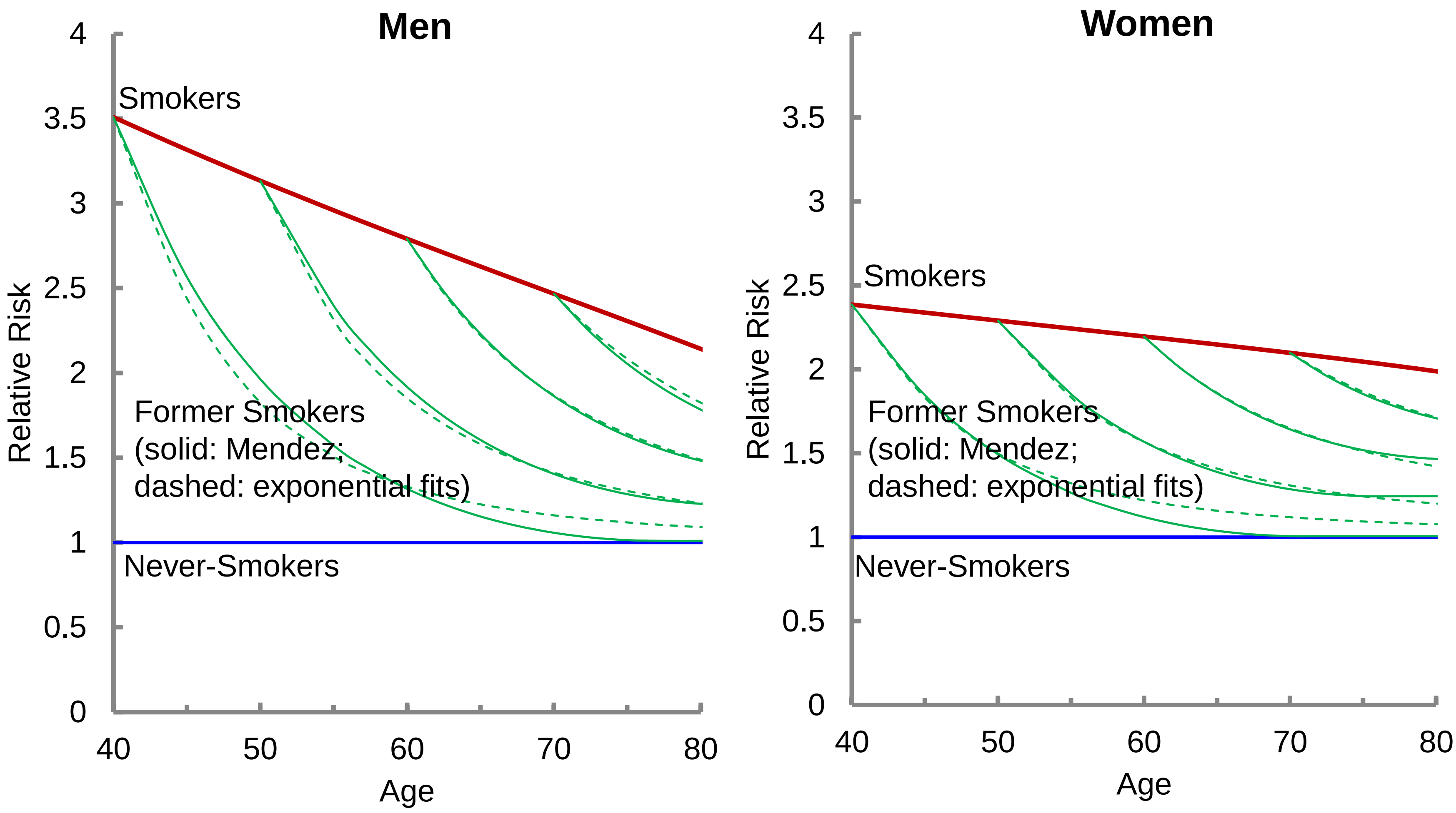
A simple relationship was fitted between CPD and Relative Risk (or Excess Risk: Relative Risk - 1).



Assumption: this curve applies to *reducers* as well as smokers with consistent CPD.
Thus, if a smoker with ER = 2.1 at 20 CPD reduces to 16 CPD, while adding an MRTP with 10% ERR, ER becomes ln(1 + 0.35*16) + 10%(2.1) = 2.1: no net change.

Poland B, Teischinger F. Population Modeling of Modified Risk Tobacco Products Accounting for Smoking Reduction and Gradual Transitions of Relative Risk. Nicotine Tob Res. 2017 Nov 1;19(11):1277-1283. doi: 10.1093/ntr/ntx070.

We modeled RR as declining exponentially after quitting, as a function of age and sex.



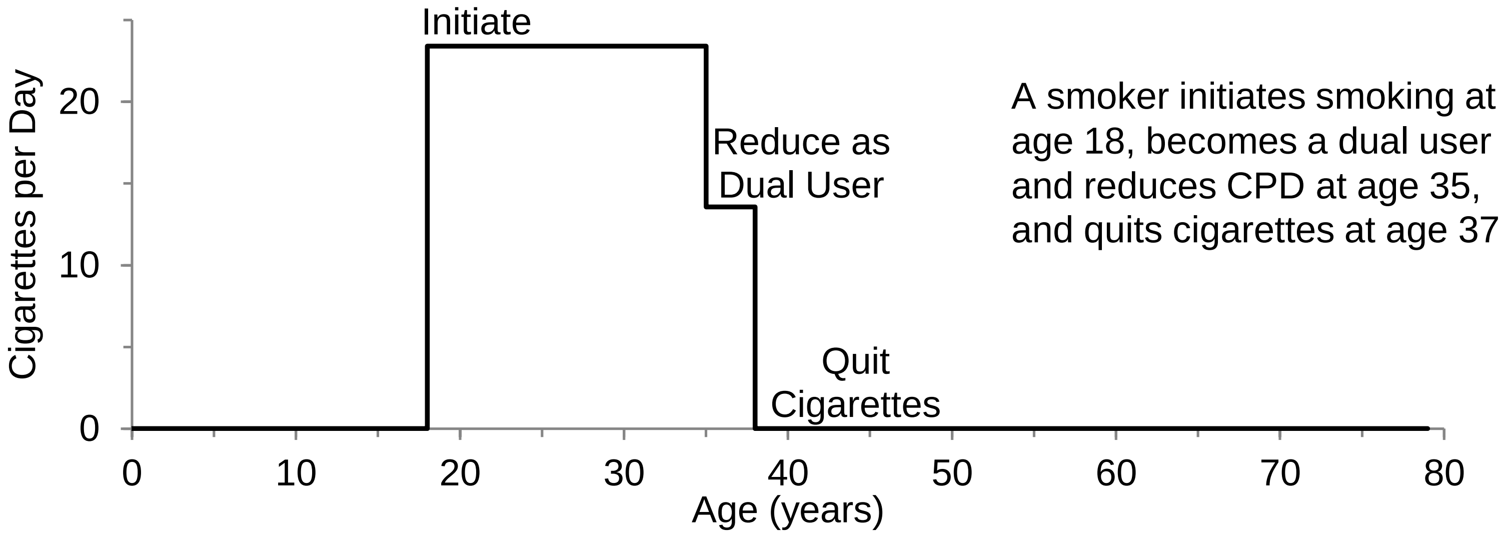
Mendez D, Warner KE 2001. The relative risk of death for former smokers: the influence of age and years-quit. Unpublished research monograph. www.umich.edu/~dmendez/tobacco/RRiskmonograph.doc

We then applied the same gradual RR change with exponential slopes to any change in RR (or ER).

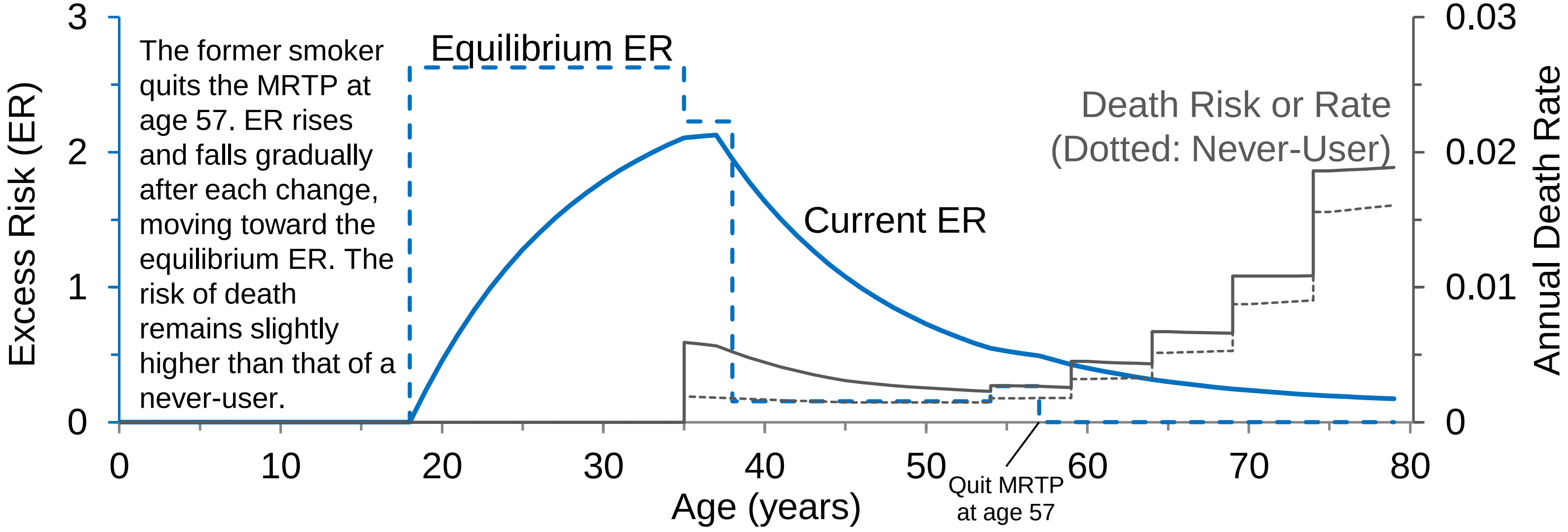
- Excess Risk $ER(t) = ER(0) \exp(-k\,t) + ER_{eq}\,[1 - \exp(-k\,t)]$
  $= ER(0) + [ER_{eq} - ER(0)]\,[1 - \exp(-k\,t)]$
  where ER_{eq} is equilibrium Excess Risk, and k is the exponential slope (a function of age and sex).
  - For a former smoker, $ER_{eq} = 0$, so $ER(t) = ER(0) \exp(-k\,t)$
  - For a new smoker, $ER(0) = 0$, so $ER(t) = ER_{eq}\,[1 - \exp(-k\,t)]$
- Conveniently, if a simulation steps through time in 1-year intervals and k is in per-year units, $ER(t)$ can be updated annually simply by setting $t = 1$ above, with the current annual values of k and ER_{eq}.
  - Think of t as relative to the current time in the simulation, so that the initial value ER(0) is the current value.
  - This provides a simple way to **account for a full tobacco product use history**, no matter how complex.

Poland B, Teischinger F. Population Modeling of Modified Risk Tobacco Products Accounting for Smoking Reduction and Gradual Transitions of Relative Risk. Nicotine Tob Res. 2017 Nov 1;19(11):1277-1283. doi: 10.1093/ntr/ntx070.

Thus, Excess Risk always moves towards equilibrium, which is also changing.



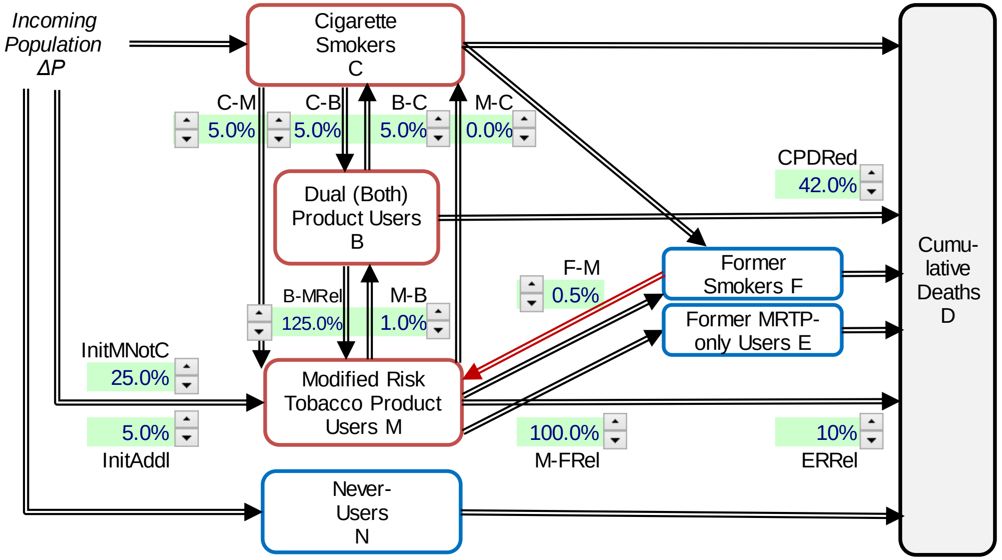
A smoker initiates smoking at age 18, becomes a dual user and reduces CPD at age 35, and quits cigarettes at age 37.



The former smoker quits the MRTP at age 57. ER rises and falls gradually after each change, moving toward the equilibrium ER. The risk of death remains slightly higher than that of a never-user.

A dozen MRTP input parameters each have Low, Base, and High cases, for sensitivity and probabilistic analysis.

Low	Base	High	
0%	25%	50%	1. InitMNotC: Would-be cigarette initiators who instead initiate MRTP (% of cigarette initiators in no-MRTP scenario)
0%	5%	10%	2. InitAddl: MRTP initiators--additional beyond would-be cig.-initiators (% of cig. initiators in no-MRTP scenario)
0.0%	0.5%	1.0%	3. F-M: Rate former smokers initiate (relapse to) MRTP (% of former cigarettes-only smokers/yr)
0%	5%	10%	4. C-B: Rate cigarette-only smokers add MRTP (% of cig-only smokers/yr)
0%	1%	2%	5. M-B: Rate MRTP (only) users add cigarettes (% of MRTP-only users/yr)
0%	42%	84%	6. CPDRed: Reduction in cig. use level by dual users (% of level without MRTP)
50%	100%	200%	7. M-F: Quit rate for MRTP (only) users (% of cigarette quit rate)
0%	5%	10%	8. B-C: Rate dual users quit MRTP (continuing cigarettes) (% of dual users/yr)
100%	125%	150%	9. B-M: Cigarette quit rate for dual users (% of cigarette-only quit rate)
5%	10%	20%	10. ERR: Proportion of cig. smoker Excess Risk experienced by MRTP users
2.5%	5.0%	10.0%	11. C-M: Rate cigarette-only smokers switch completely to MRTP (%/yr)
	0%	2%	12. M-C: Rate MRTP-only users switch completely to cigarettes (%/yr)



Simulate 1,000,000 individuals through year 2100

☑ Also simulate no MRTP

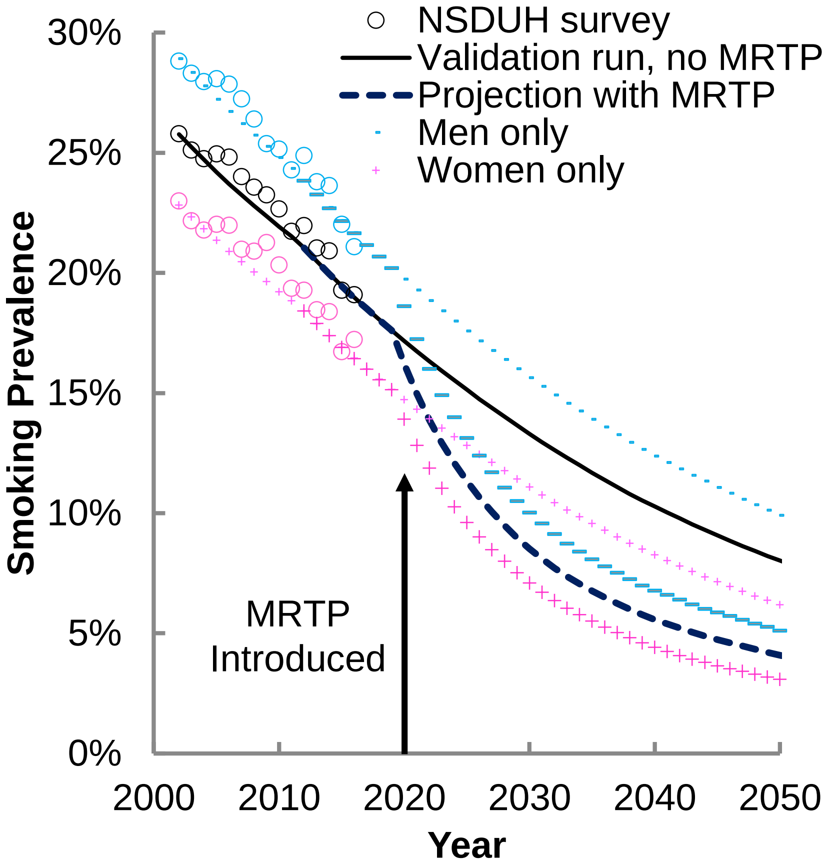
☐ Also simulate low & high ERR

By also simulating a no-MRTP scenario, we can calculate Avoided Premature Deaths.

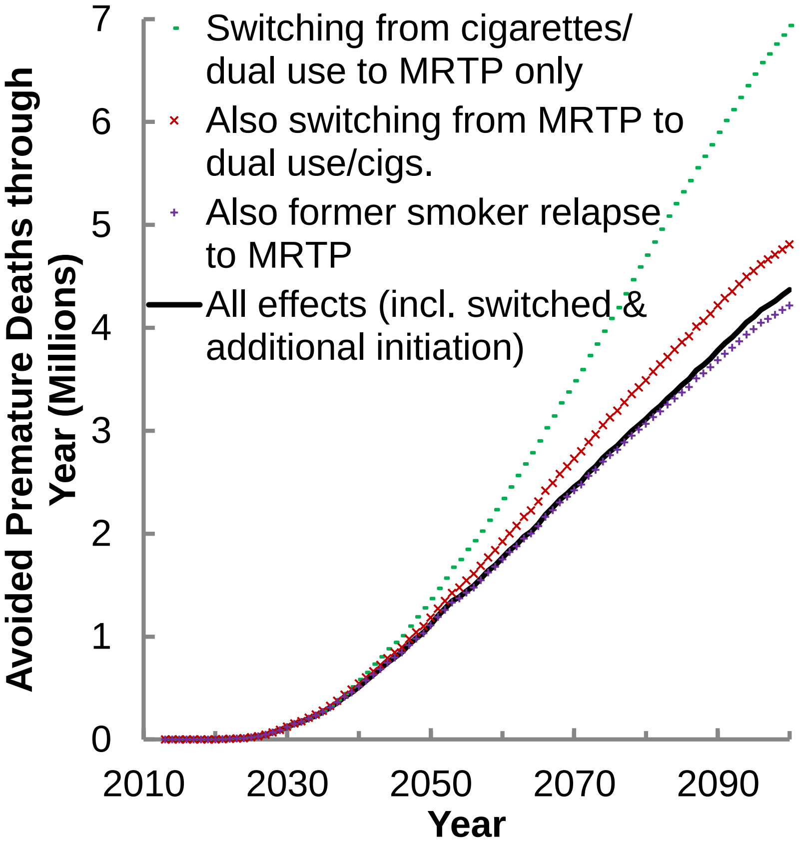
The model was validated with NSDUH prevalence data and used to project avoided premature deaths.

Avoided premature deaths = deaths occurring later *with* MRTP
minus deaths occurring later *without* MRTP.

Comparison with data

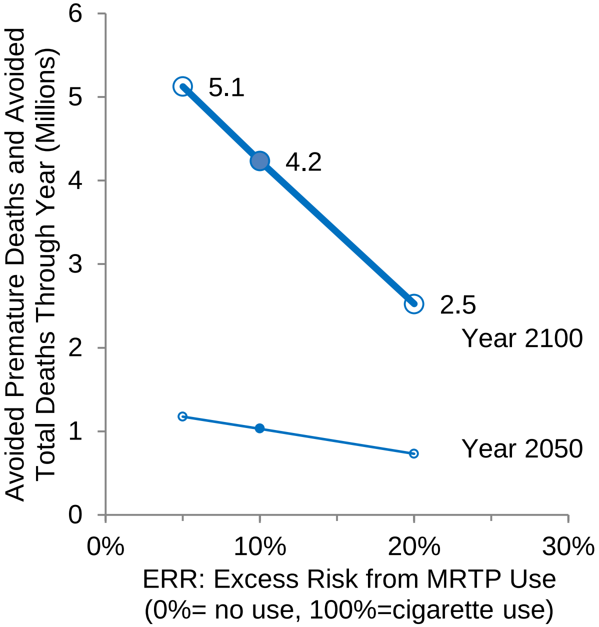


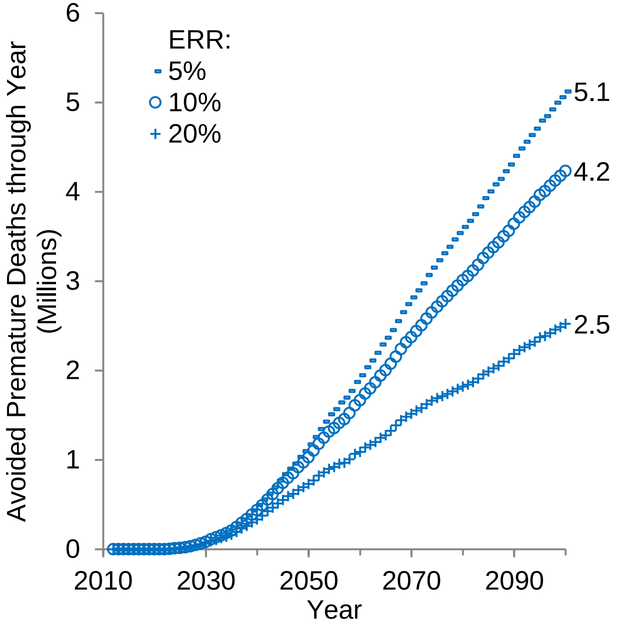
Avoided premature death breakdown

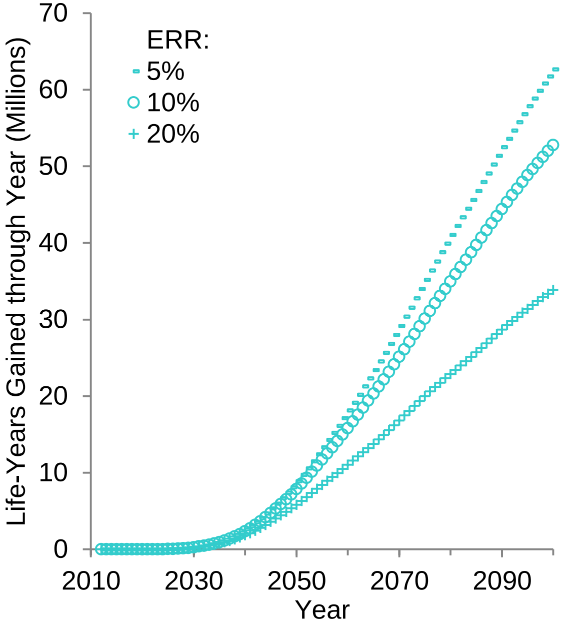


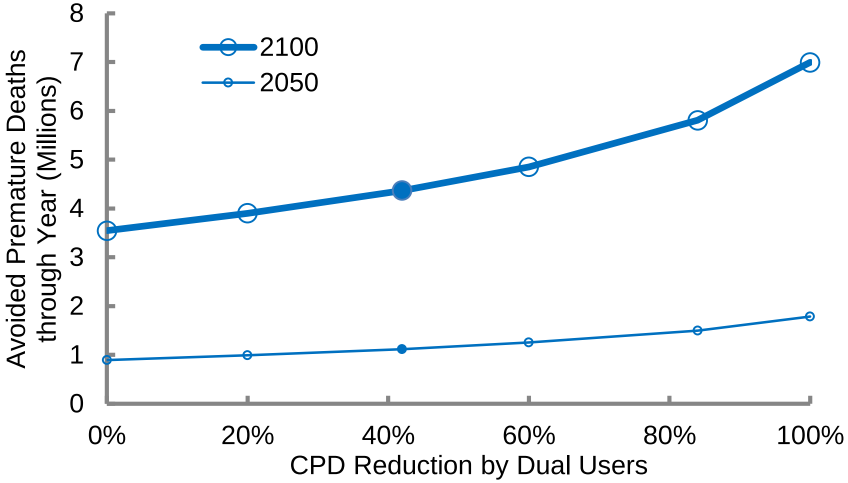
Avoided total (cumulative) deaths don't work well in later years because those whose deaths are avoided initially die eventually.

Outputs include sensitivities to ERR (remarkably linear, allowing quick approximation of break-even) and to CPD, as well as life-years gained.

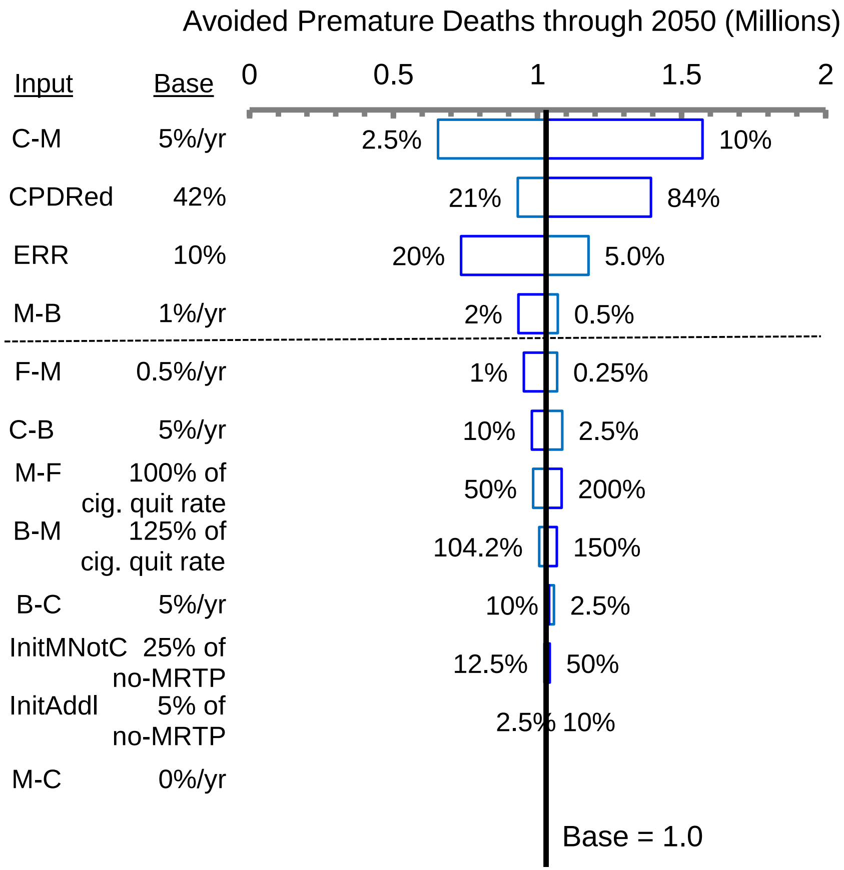


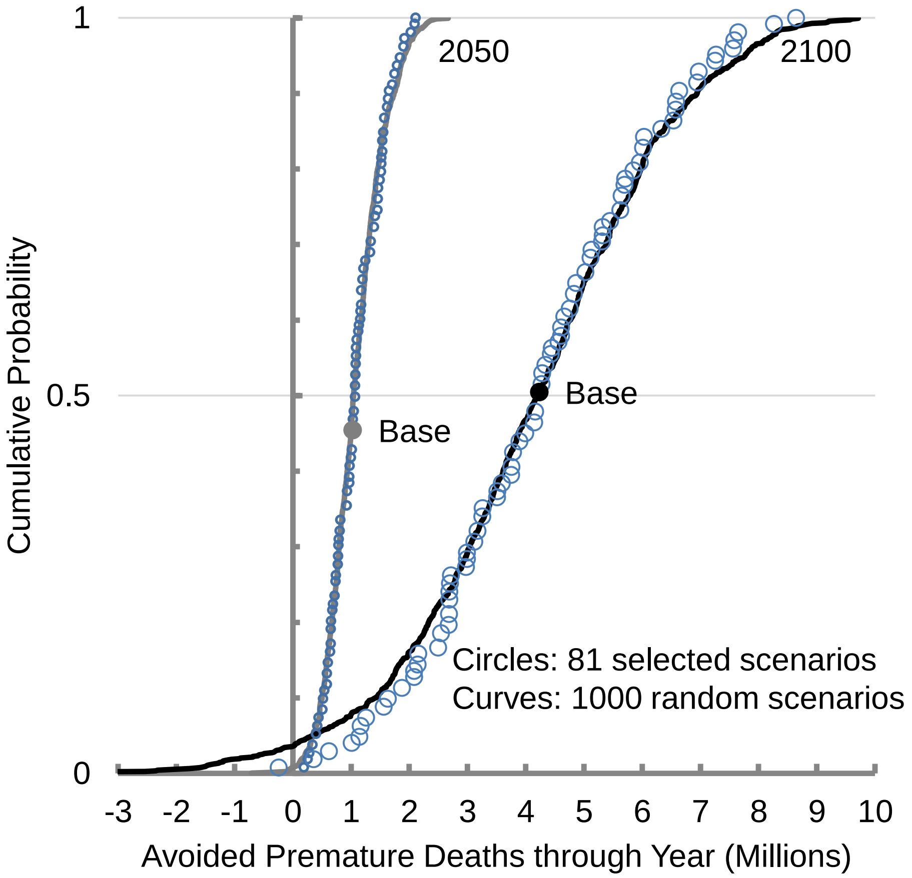






A tornado chart highlights the most sensitive inputs. Full and simplified probabilistic analysis methods show good agreement.





Bars show effect of varying each input one-at-a-time between low and high cases; base line shows output with all inputs at base values.

Curves use 1000 Monte Carlo draws from fitted lognormals for **all 12** inputs. Circles use a tree of 3^4 = 81 combinations of low (p=0.3), base (p=0.4), and high (p=0.3) values of **top 4** tornado inputs.

Poland B, Larroque S. Probabilistic Analysis of the Effects of a Modified Risk Tobacco Product on Population Health. SRNT Annual Meeting, Baltimore, MD, Feb. 21-24, 2018, Abstract POS4-134. https://www.jt-science.com/sites/default/files/2018-04/2018-p1-jti.pdf

The Monte Carlo probabilistic analysis is simply an outer loop around the model, with different random sets of parameters.

- The base model itself happens to be Monte Carlo, but it could have been a system dynamics or other model.

- Run time is an issue if the model is slow:
  - 1000 probabilistic runs x 1 million people each = 1 billion people simulated!

- Theoretical question (when the base model also uses Monte Carlo): how best to balance the size of outer and inner loops.
  - Maybe we don’t need 1 million people.

The probability tree analysis uses m^n (here $3^4 = 81$) runs, combining *m* cases for the top *n* variables from the tornado.

- Here we use the low, base, and high cases for C-M, CPDRed, ERR, and M-B.
- Thus we can easily label and interpret each scenario.
    - For example, the worst scenario (at bottom left of cumulative curve) has low C-M, low CPDRed, high ERR, and high M-B.
- Assuming the low, base and high cases represent 10th, 50th, and 90th percentiles, we assign them probabilities 0.3, 0.4, and 0.3.
    - Thus the 81 scenarios have probabilities ranging from 0.3^4 to 0.4^4 (for the base case), summing to 1.
    - These probabilities preserve the mean and variance of a *normal* distribution with these percentiles.

Each approach has advantages. A hybrid approach adding Monte Carlo scenarios to the tree can provide a compromise.

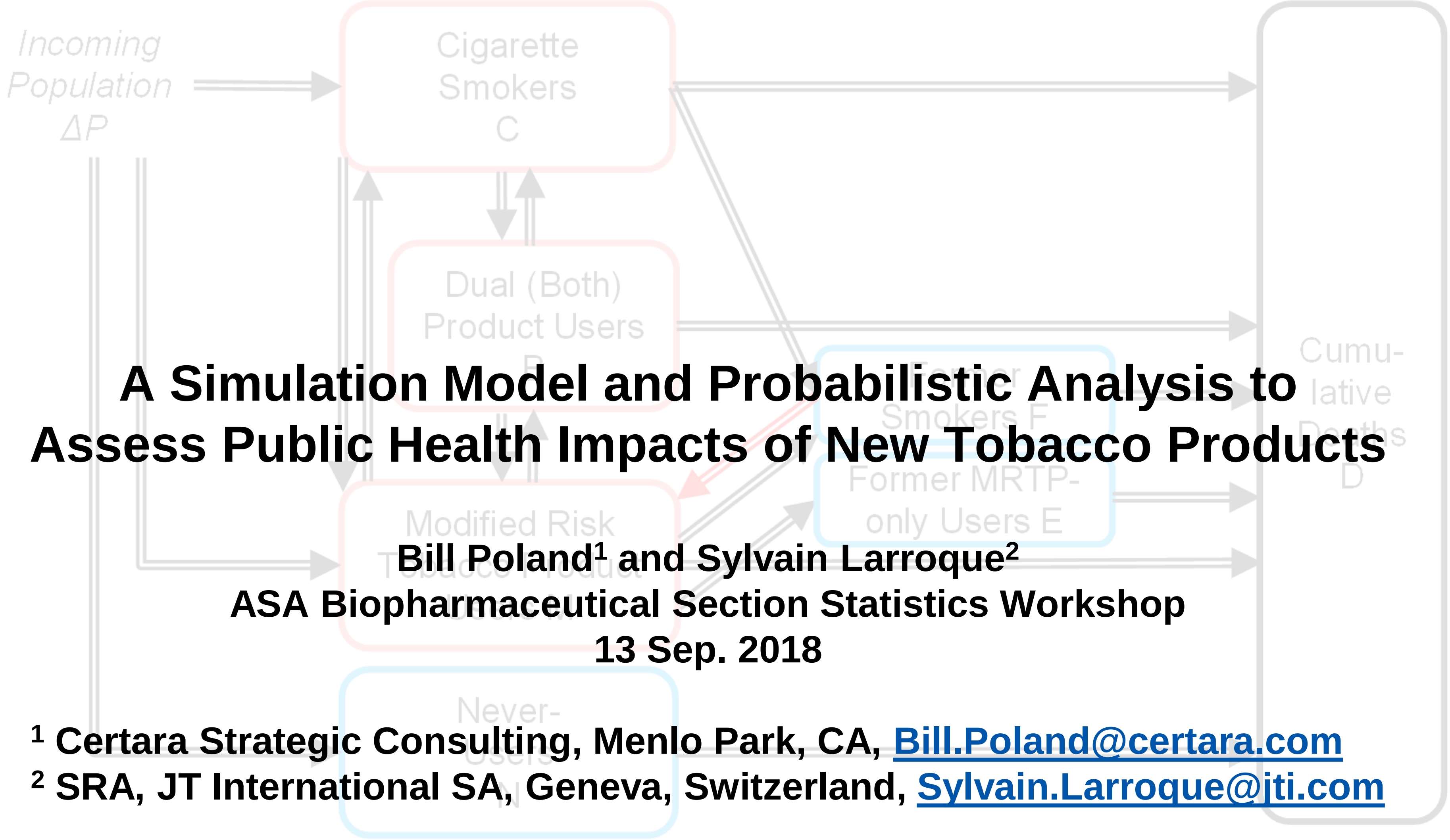
Probability Tree

- Accuracy is adjustable (*m* and *n*) and can be better than Monte Carlo for a given number (e.g., 81) runs
  - assuming most of the variance has been captured with the selected variables (rapidly narrowing tornado).
- Scenarios are easy to list and interpret, as combinations of low, base, and high cases.
- Probability trees are natural for modeling dependence among variables.

Monte Carlo

- Accuracy is adjustable simply through the number of runs.
  - Quasi-random (e.g., Sobol) numbers can be used to accelerate convergence.
- Easy to explain and implement.
  - Runs are weighted equally.
- Any number of variables can be treated as probabilistic with almost no extra computation.
  - No need to select key variables, so more foolproof.

Hybrid example: adding 81 random scenarios to the 81 tree scenarios improved accuracy in the tails moderately. Research question: how to optimally balance the number of each type in this hybrid approach.



A Simulation Model and Probabilistic Analysis to Assess Public Health Impacts of New Tobacco Products

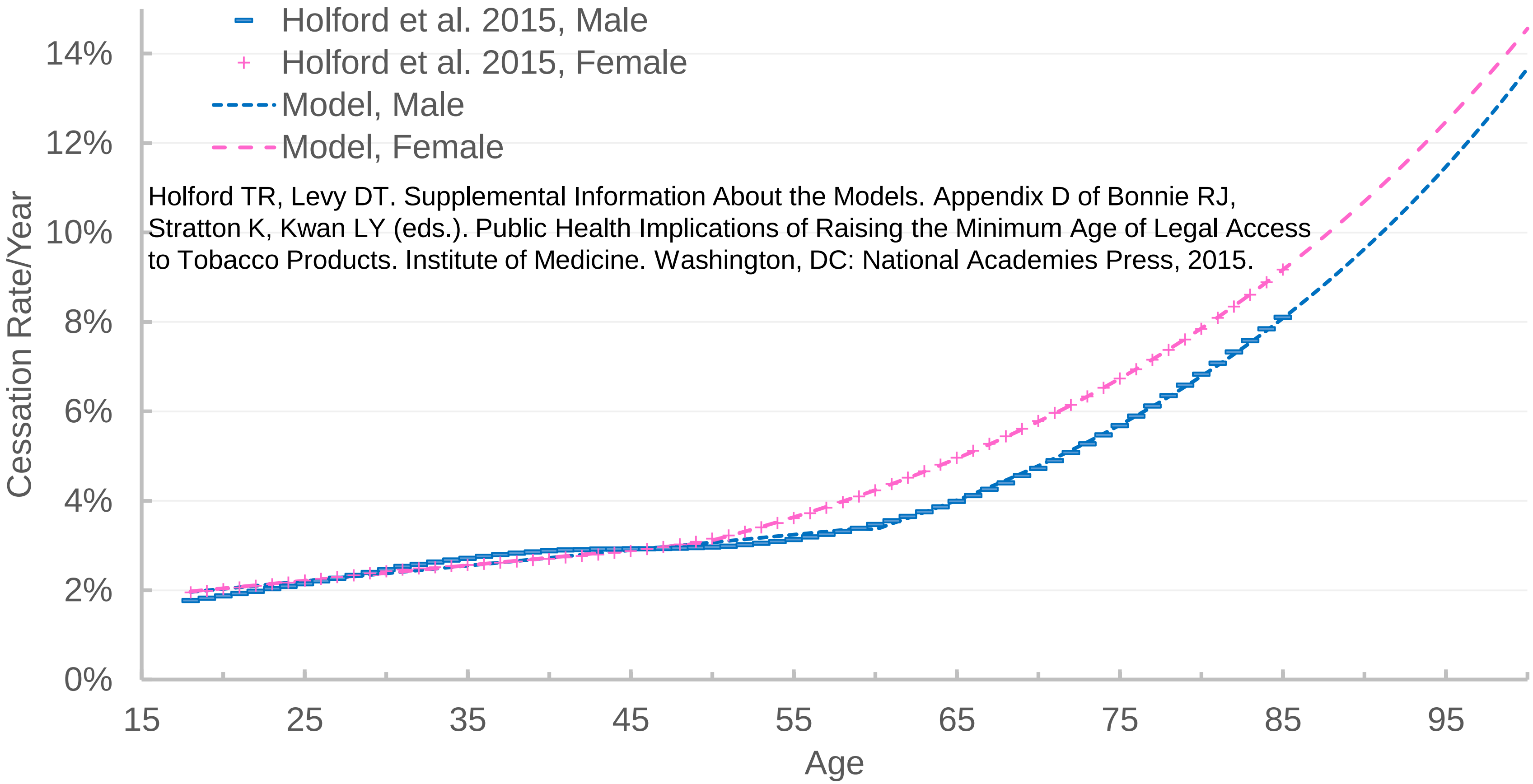
Bill Poland[1] and Sylvain Larroque[2]

ASA Biopharmaceutical Section Statistics Workshop

13 Sep. 2018

[1] **Certara Strategic Consulting, Menlo Park, CA, Bill.Poland@certara.com**
[2] **SRA, JT International SA, Geneva, Switzerland, Sylvain.Larroque@jti.com**

BACKUP. Cigarette cessation rate curves by age and sex were based on Holford et al. 2015. Initiation rates were based on NSDUH.

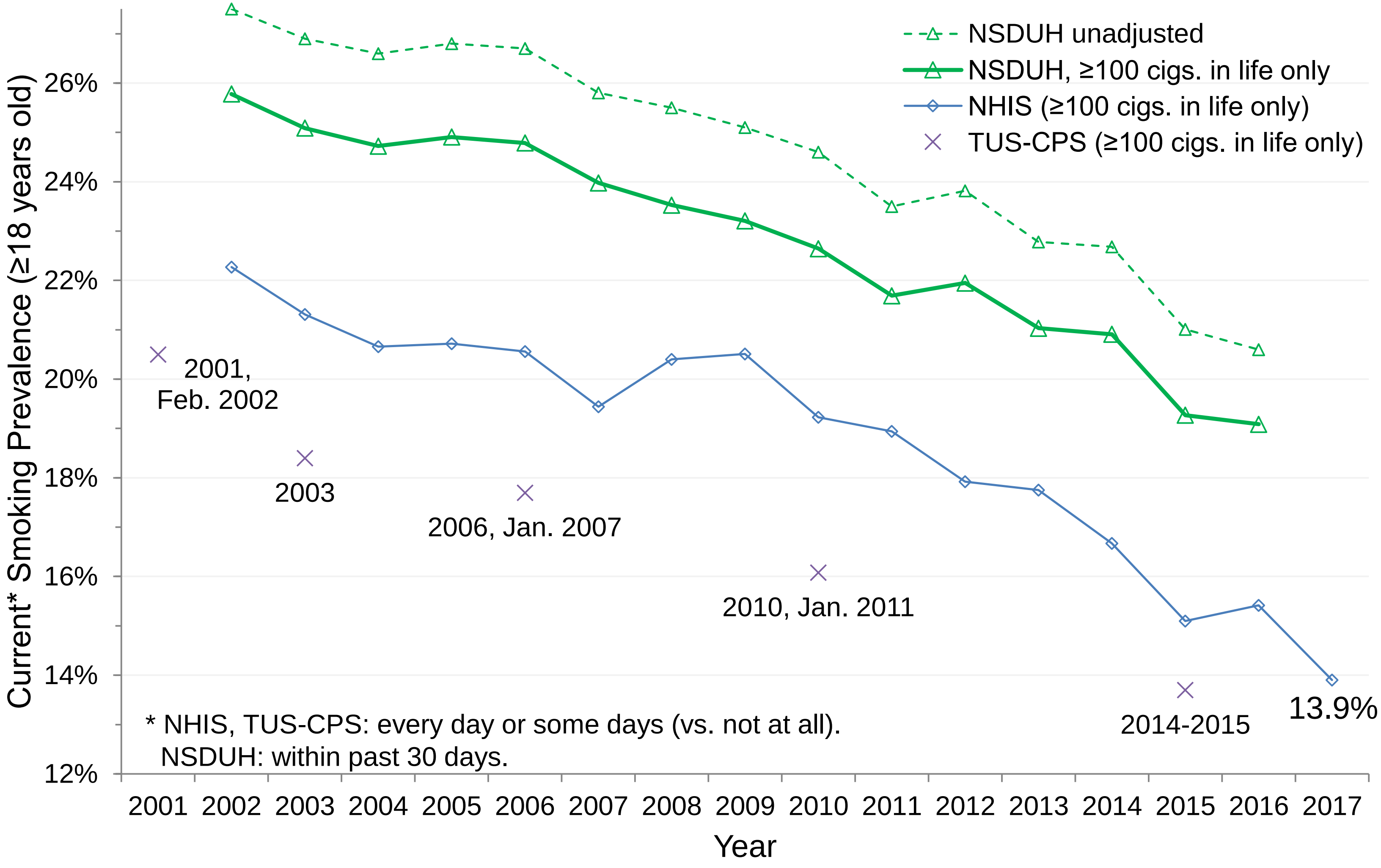


Holford TR, Levy DT. Supplemental Information About the Models. Appendix D of Bonnie RJ, Stratton K, Kwan LY (eds.). Public Health Implications of Raising the Minimum Age of Legal Access to Tobacco Products. Institute of Medicine. Washington, DC: National Academies Press, 2015.

Overall mean: 3.5%. Increase to 4.2%-4.5% based on Mendez et al. 2016?

Cigarette initiation rates were based on maximum NSDUH smoking prevalence through age 25 in the first model year (2012) by sex, with an annual decline fitted to NSDUH data since 2002.

US smoking prevalence may be declining a little faster in the e-cigarette era (last 5+ years). Earlier long-term projections now seem pessimistic.



* NHIS, TUS-CPS: every day or some days (vs. not at all).
NSDUH: within past 30 days.

Is US smoking prevalence already lowered by e-cigarettes?

We can use the two components of Excess Risk, from cigarettes and the MRTP, to break down the probability of death by cause.

- E.g., lung cancer/other cancer/CV/respiratory/other
  - Smoker and never-smoker frequencies are in Thun et al. 2013, 1997.
- For each simulated individual i and cause of death j:

$$\mathbf{RelP_{ij} = w_j\ (1 + ER_{ij}\ )\ /\ (1 + ERc_i + ERm_i)}$$

$$ER_{ij} = (ERc_{0j}\ /\ ERc_0)\ ERc_i + (ERm_{0j}\ /\ ERm_0)\ ERm_i$$

$$ERc_0 = \sum w_j\ ERc_{0j}, \qquad ERm_0 = \sum w_j\ ERm_{0j}$$

$$ERm_{0j} = ERR_j\ ERc_{0j}$$

where

  - w_j is the *never-smoker* relative probability of death by cause,
  - ERc_i and ERm_i are the individual Excess Risks for cigarettes and for the MRTP (from the model),
  - ERc_{0j} and ERm_{0j} are the Excess Risks by cause of death for current smokers and for MRTP users (varying by age group and sex also),
  - ERc_0 and ERm_0 are the corresponding overall Excess Risks, and
  - ERR_j is the MRTP Excess Relative Risk by cause (additional input).
- We can then average these by age group, sex, and time period.

Thun MJ, Carter BD, Feskanich D, et al. 2013. 50-year trends in smoking-related mortality in the United States. *N Engl J Med.* 2013;368(4):351–64. http://www.nejm.org/doi/pdf/10.1056/NEJMsa1211127, http://www.nejm.org/doi/suppl/10.1056/NEJMsa1211127/suppl_file/nejmsa1211127_appendix.pdf

Thun MJ, Myers DG, Day-Lally C, et al. 1997. Age and the exposure-response relationships between cigarette smoking and premature death in Cancer Prevention Study II. Chapter 5 in: Changes in Cigarette-Related Disease Risks and Their Implications for Prevention and Control. Smoking and Tobacco Control Monograph No. 8. Burns DM, Garfinkel L, Samet JM, eds. Bethesda (MD): U.S. DHHS 1997:383–475. NIH Publication No. 97-4213. http://cancercontrol.cancer.gov/brp/TCRB/monographs/8/m8_complete.pdf