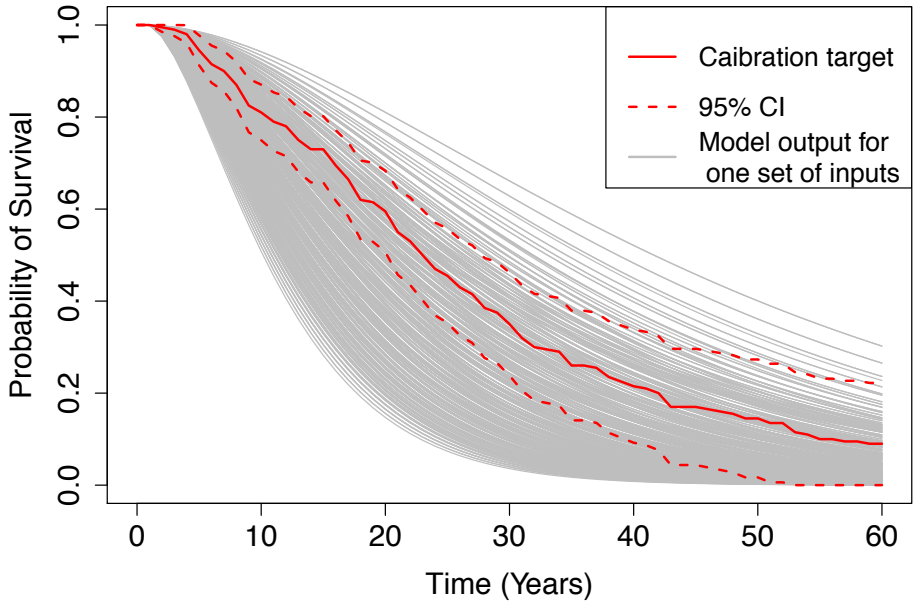
Introduction to model calibration

Motivation

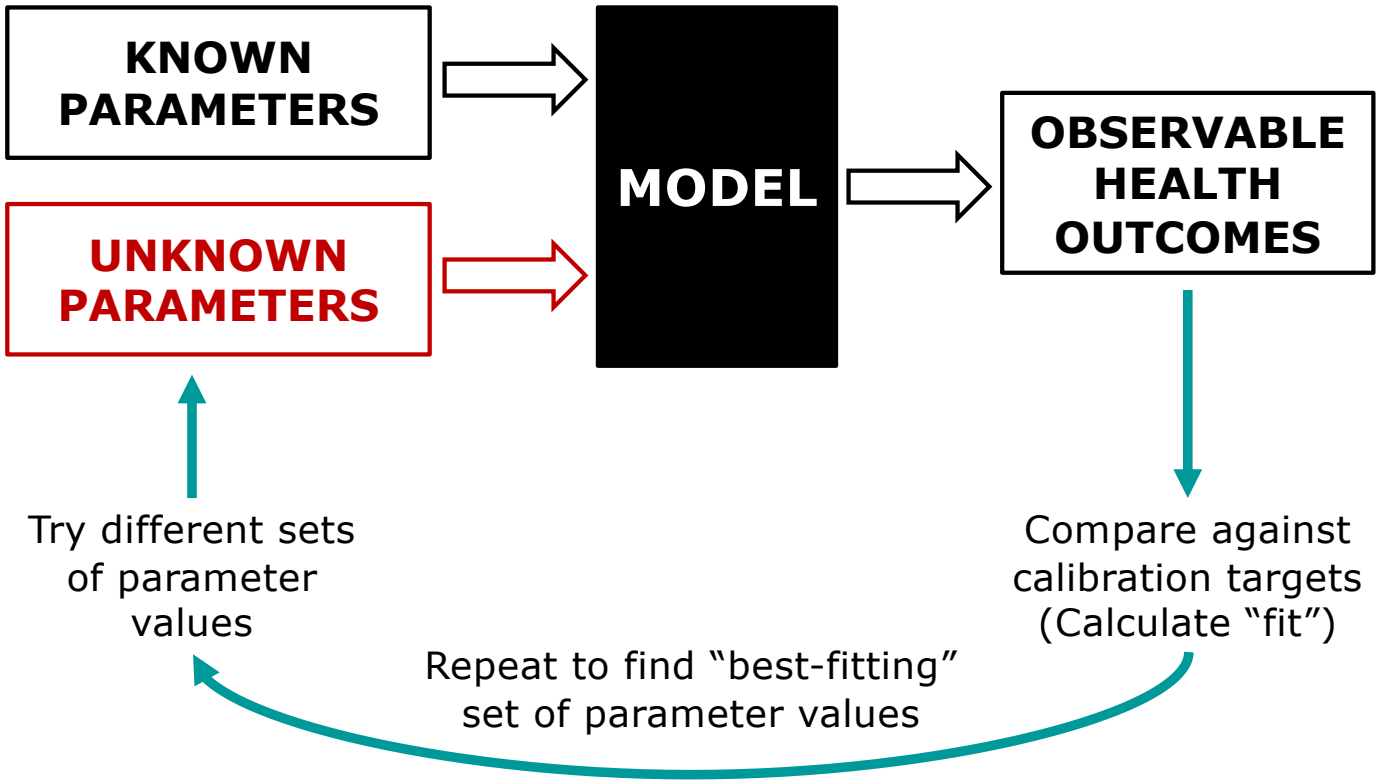
- Mathematical models of disease often involve a subset of parameters whose values are unknown
- Common reasons include physical, feasibility, and ethical limitations
- Estimate values for these parameters by matching model outputs to observed outcomes
  - Model calibration

Calibration definition

- Process of adjusting model input parameter values to match data on an outcome of interest (e.g., survival, prevalence, or incidence)
- Outcome(s) of interest = "calibration targets"

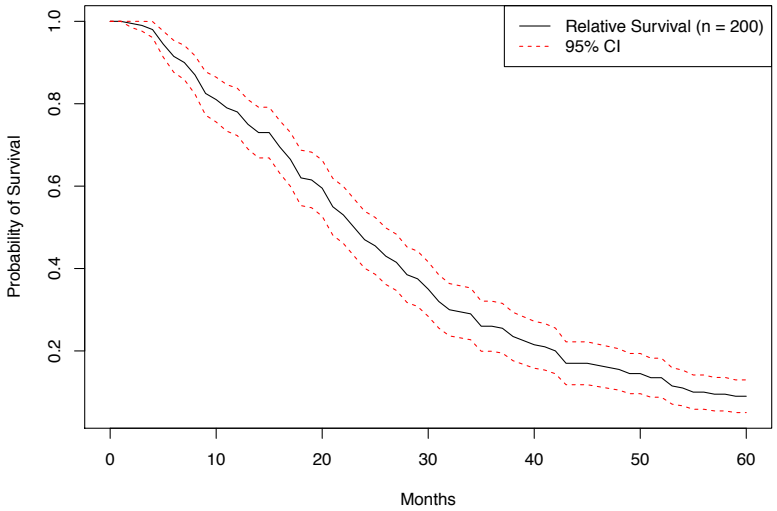


Calibration process



Calibration targets

- Empirical data to be replicated by the model
- Summary statistics (e.g. mean age of cancer diagnosis) or series of observations (e.g. age-specific incidence)
- Can calibrate to multiple targets (e.g. survival and prevalence) simultaneously



Calculating “fit”

- Goodness-of-Fit (GoF) is the quantitative measure of how the model is replicating the target data
- Different ways to measure GoF
  - Distance
  - Likelihood

Distance GoF measures

- Notation
  - M : a mathematical model (e.g., Markov model)
  - θ : Set of K parameters to be calibrated
  - y : Values of T calibration targets

- Sum of squared errors

$$SSE(\theta) = \sum_{i=1}^{T} \left(y_i - M_i(\theta)\right)^2$$

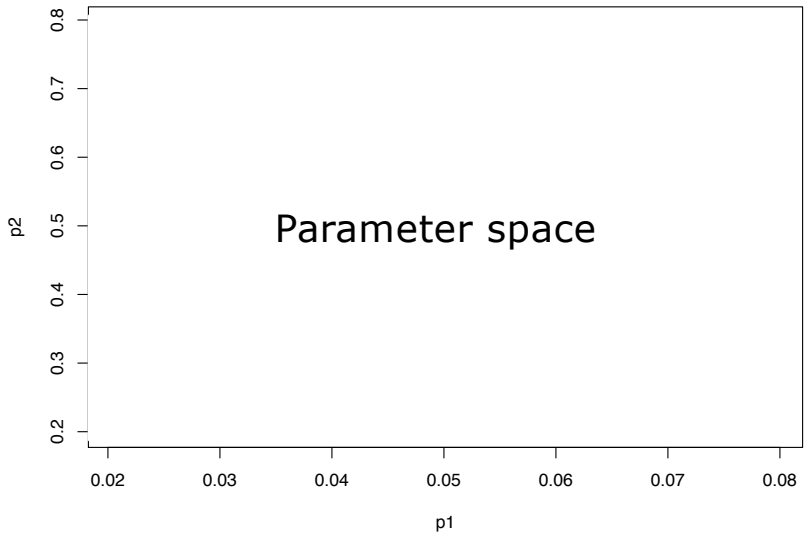
Weighted sum of squared errors

$$WSSE(\theta) = \sum_{i=1}^{T} w_i \left(y_i - M_i(\theta)\right)^2$$

(Often $w_i = \frac{1}{\sigma^2}$)

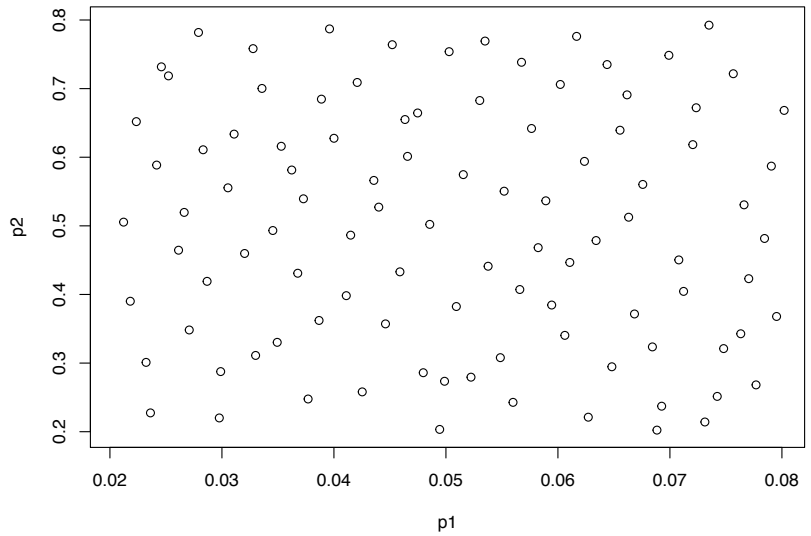
Search Strategy

- Define plausible ranges for parameter whose values are unknown
- Use a search strategy to "search" through the input parameter space
  - Run the model for sets of parameter values generated by search strategy



Random Search

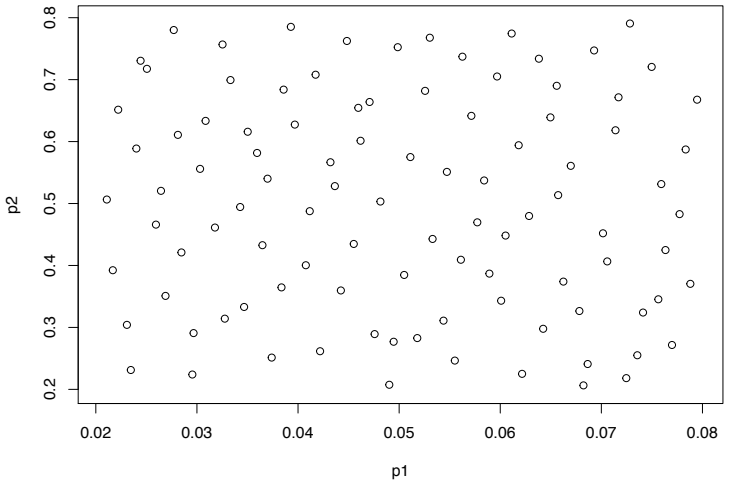
- Randomly sample a large number of parameter value sets from probabilistic distributions
- Use "Latin hypercube sampling" (LHS) to ensure sample captures the full parameter space



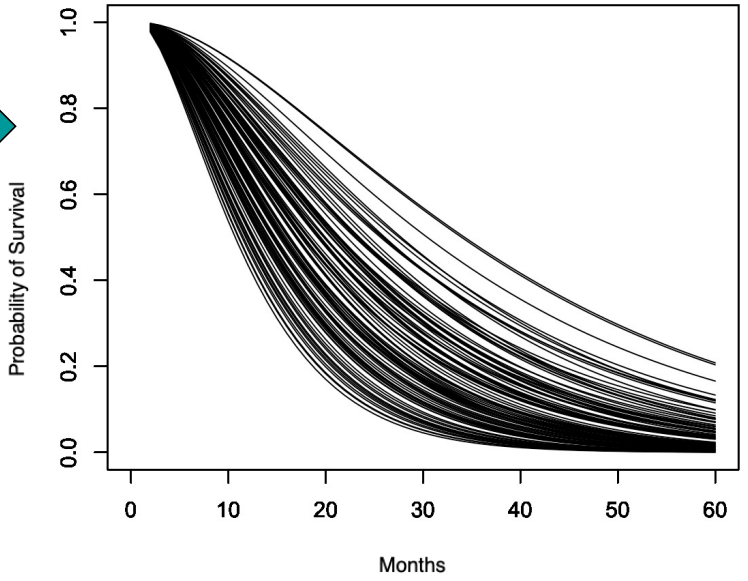
Calibration Illustration --

- Target: survival curve
- GOF: Weighted sum of squared errors
- Search strategy: Random search

Generate random sample of sets of parameter values from parameter space



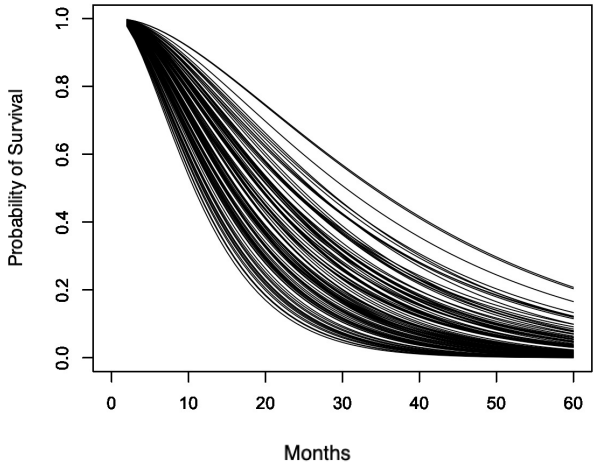
For each set, run model and generate output corresponding to calibration target



Calibration Illustration --

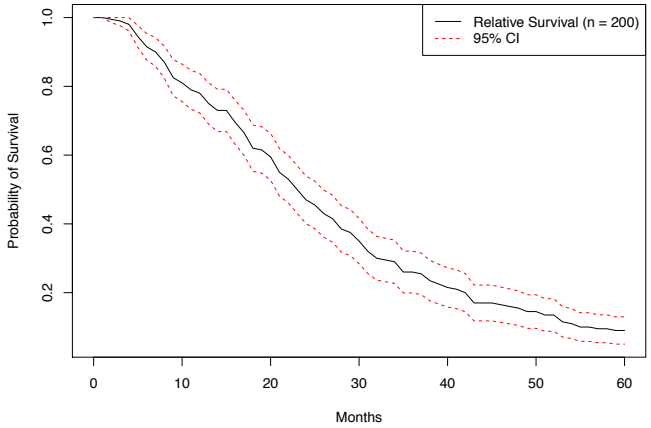
- Target: survival curve
- GOF: Weighted sum of squared errors
- Search strategy: Random search

Compare each output to targets and calculate WSSE



Model outputs

vs.



Target data

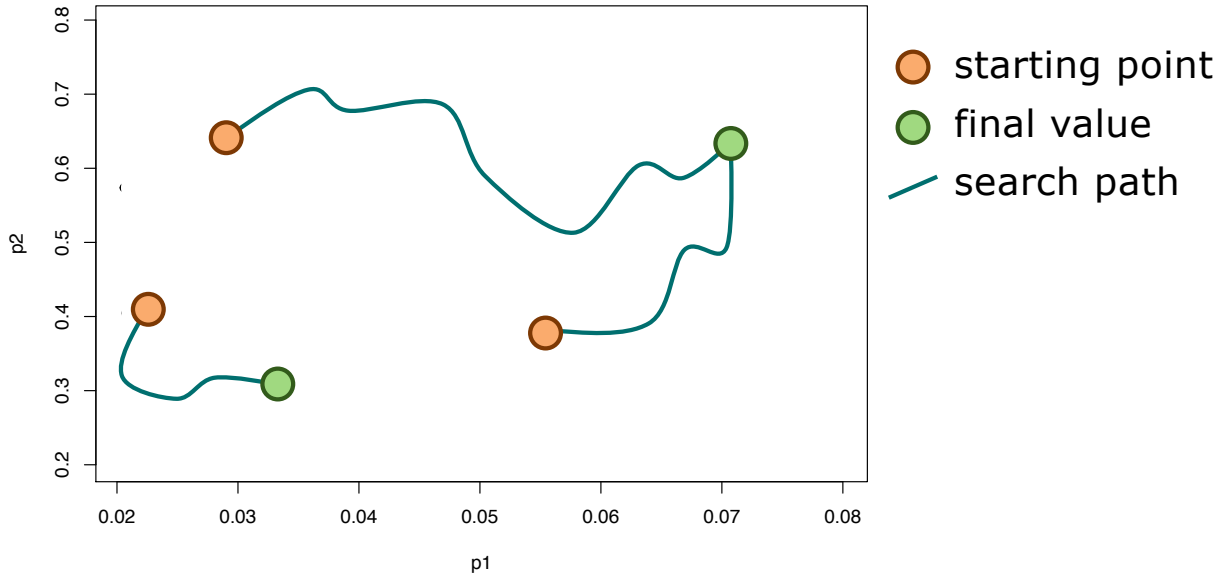
p1	p2	fit
0.26354529	0.06179677	1413.7403
0.09051233	0.07470821	3154.9646
0.30208885	0.01378525	2176.8862
0.73138604	0.06422640	1698.5179
0.78408031	0.05221322	722.9070
0.73380800	0.07237515	663.1347
0.50475719	0.07211332	149.9460
0.01438836	0.03284466	1189.5260
0.18840177	0.03172401	940.5997
0.41323042	0.06268992	136.3939

Iterative Search

- Use fits of past input values to determine which input values to try next
- Directed methods
  - Nelder-Mead (simplex method)
  - Gradient–descent and others
- Meta-heuristic algorithms
  - Genetic algorithms
  - Simulated annealing

Nelder-Mead Algorithm

- Downhill simplex method
- Must be run multiple times for different starting points to avoid local extrema



DARTH Workgroup

Fernando Alarid-Escudero, PhD[1]
Eva A. Enns, MS, PhD[2]
M.G. Myriam Hunink, MD, PhD[3,4]
Hawre J. Jalal, MD, PhD[5]
Eline M. Krijkamp, MSc[3]
Petros Pechlivanoglou, PhD[6]
Alan Yang, BSc[6]

In collaboration of:
1 Division of Public Administration, Center for Research and Teaching in Economics (CIDE), Aguascalientes, Mexico
2 University of Minnesota School of Public Health, Minneapolis, MN, USA
3 Erasmus MC, Rotterdam, The Netherlands
4 Harvard T.H. Chan School of Public Health, Boston, USA
5 University of Pittsburgh Graduate School of Public Health, Pittsburgh, PA, USA
6 The Hospital for Sick Children, Toronto and University of Toronto, Toronto ON, Canada

www.darthworkgroup.com

CIDE

SCHOOL OF PUBLIC HEALTH
UNIVERSITY OF MINNESOTA

Erasmus MC
Netherlands Institute for Health Sciences

University of Pittsburgh
Graduate School of Public Health

SickKids