Modeling Cumulative Evidence for Freedom from Disease with Applications to BSE Surveillance Trials

Dankmar Böhning

Applied Statistics
School of Biological Sciences
University of Reading

Presentation at:

Statistics Research Seminar Series
School of Mathematics
University of Southampton

May 04, 2006
Organizer: Dave Woods

- Joint work with Matthias Greiner (Head: Intenational EpiLab, Danish Institute for Food and Veterinary Research, DK)
- Work done while visiting International EpiLab in 2004

- Published details available:
  - Böhning & Greiner (2006, EUJE) *(study part)*
  - Böhning & Greiner (2006, JABES) *(theoretical part)*

www.reading.ac.uk/~sns05dab/

Overview

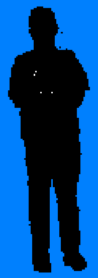
- Background
- Idea and Scope of the project
- Preliminary Results
- Situation in Denmark
- Non-Perfect Diagnostic Testing
- Incorporating Heterogeneity from Different Surveillance Streams

Overview

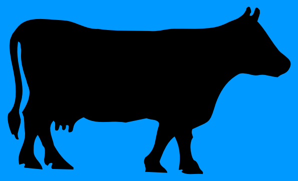
- Background
- Idea and Scope of the project
- Preliminary Results
- Situation in Denmark
- Non-Perfect Diagnostic Testing
- Incorporating Heterogeneity from Different Surveillance Streams

Background

- BSE risk mitigation
  - prevent specific risk materials from entering the food chain
  - restrict import of food with BSE risk
  - safe processing of food
  - detect BSE before beef enters food chain

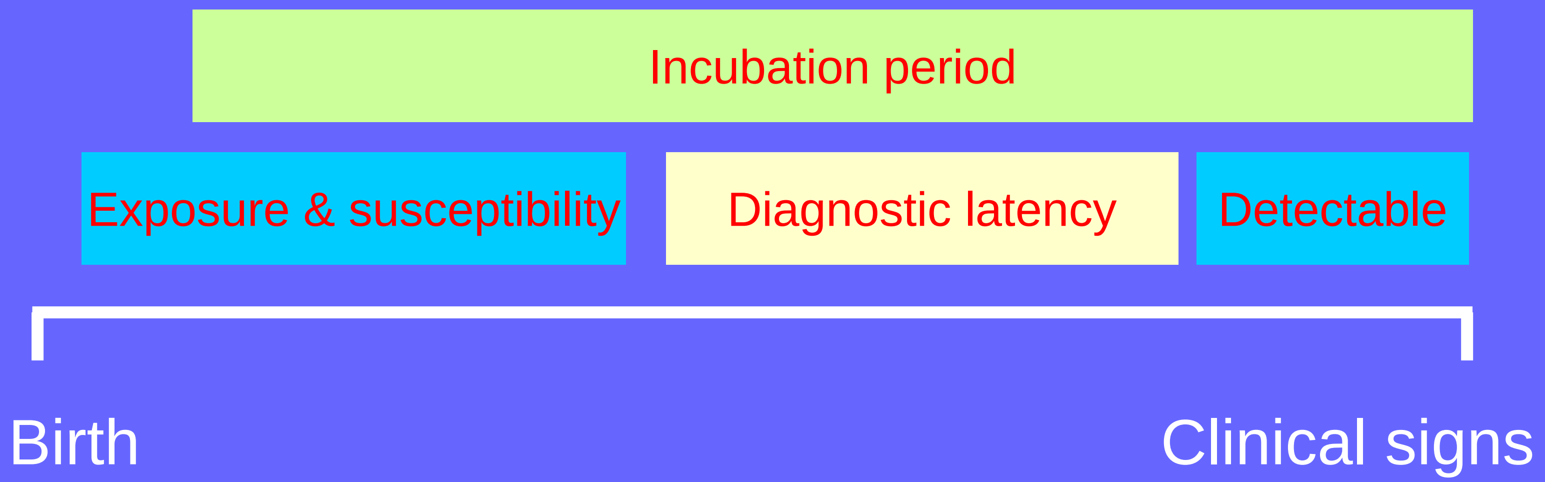


  - remove specific risk materials from animal feed
  - restrict import of cattle with BSE risk
  - ban meat-and-bone-meal for ruminants
  - destroy BSE infected bovines
  - BSE surveillance in cattle



Background

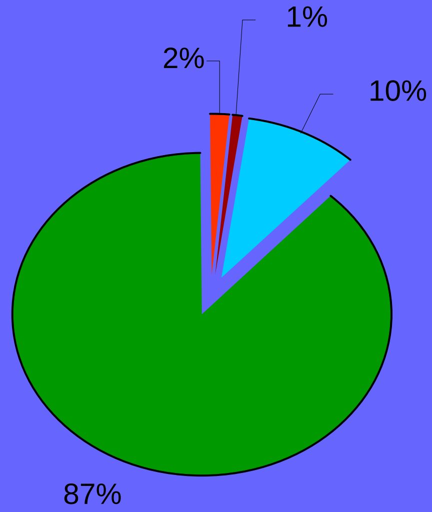
- BSE is a slow disease
  - BSE cases reflect exposure in the past
  - animals in early incubation phase cannot be diagnosed



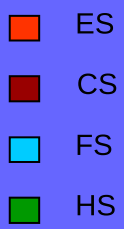
Background

- BSE surveillance in the EU
  - all fallen stock (FS) >24 months
  - all emergency slaughtered (ES) cattle >24 months
  - all healthy slaughters (HS) > (24) 30 months
  - all clinical suspects (CS) > 24 months

- EU Regulation No. 999/2001

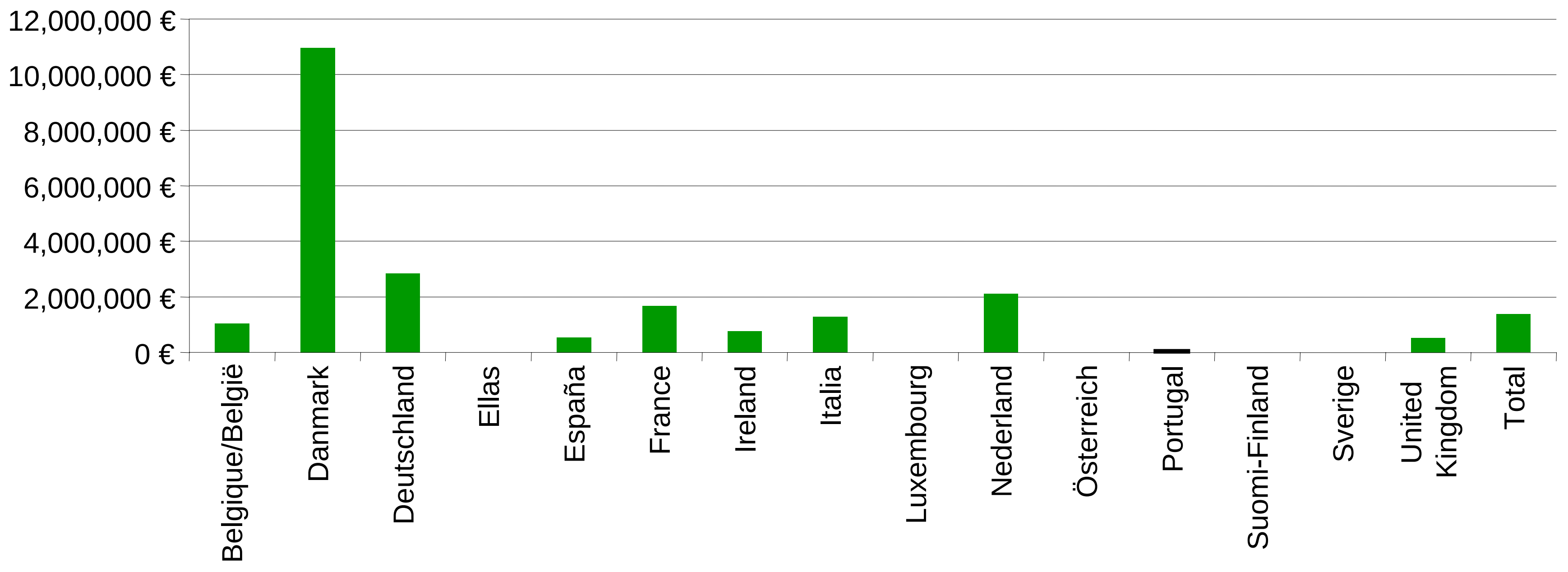


Distribution of BSE tests (EU, 2002)



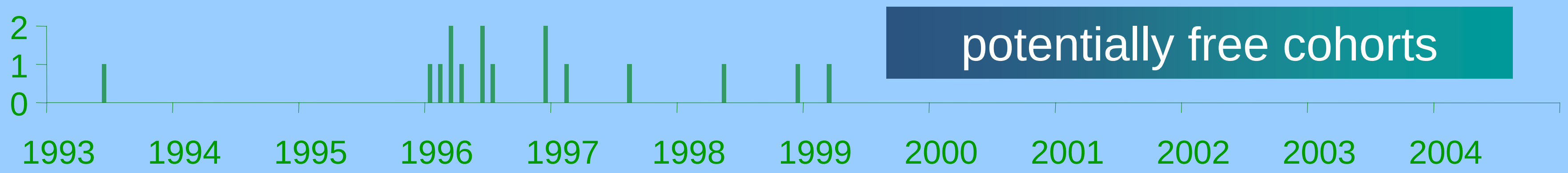
Background

Testing for BSE is expensive



Background

Number of Danish BSE cases by birth cohort (month)



Data source:

http://www.dfvf.dk/Default.asp?ID=9827

Objectives

- To develop a statistical approach suitable for documenting freedom from BSE, stratified for birth cohorts, which
  - will account for the longitudinal data flow from distinct surveillance streams
  - allow adaptive up-scaling and down-scaling of the sampling coverage and optimal allocation of testing resources to birth cohorts based on prior risk estimates
  - contribute to a critical review of the current zero prevalence policy for BSE surveillance

Overview

- Background
- Idea and Scope of the project
- Preliminary Results
- Situation in Denmark
- Non-Perfect Diagnostic Testing
- Incorporating Heterogeneity from Different Surveillance Streams

Idea of Project

- birth cohorts of animals (in different surveillance streams) are monitored for occurence of BSE
- in particular, prevalence is small, potentially cohort is disease-free
- in contrast to estimating prevalence, this project wants to answer the question:
- When can a particular cohort considered to be disease free?

π

Idea of Project
Basic Principle of the Sequential Trial

interest is in a prevalence parameter π and associated null hypothesis

$$H_0 : \pi = 0$$

(implying, birth cohort is disease-free)

π

sequential trial (ST): animals are tested in discrete calendar or sequential time

Y_t result of testing animal t

$(y_t = 1$ test positive, $y_t = 0$ test negative$)$:

$H_0: \; Y_t = 0$ for all times $t = 1, 2, 3, \ldots$

clearly, $\Pr(Y_t > 0 \mid H_0) = 0$, for all times t

in other words, there is no type-I error

$Y_1, Y_2, Y_3 ...$ series of BSE-tests:

waiting time T for first animal testing positive:

$$\Pr(T = t \mid \pi) = \pi(1-\pi)^{t-1}$$

has geometric distribution

T	sequence of tests	probability
1	1	π
2	01	$(1-\pi)\pi$
3	001	$(1-\pi)^2\pi$
4	0001	$(1-\pi)^3\pi$
....	....	.....

Rationale of the ST:

since

$$\Pr(T > 0 \mid \pi) = \sum_{t=1}^{\infty} \pi(1-\pi)^{t-1} = 1$$

unless $\pi = 0,$ there exists some positive time waiting time $s > 0$ such that

$$\Pr(0 < T \le s \mid \pi) = 1 - \beta$$

for given arbitrary small $\beta > 0$

Rationale of the ST:

instead of waiting for all times ($T=\infty$)

to conclude with $\pi = 0$,

we wait until time $s<\infty$ such that

$$\Pr(0<T\le s\,|\,\pi)=\sum_{t=1}^{s}\pi(1-\pi)^{t-1}=1-\beta$$

to conclude with $\pi = 0$, necessarily.

now,

$$\Pr(0 < T \le s \mid \pi) = \sum_{t=1}^{s} \pi(1-\pi)^{t-1} = 1-(1-\pi)^{s}$$

and equating

$$1-(1-\pi)^{s} \;\; = 1-\beta$$

leads to

$$(1-\pi)^{s} = \beta$$

Idea of Project: Solution

$$(1-\pi)^s = \beta$$

from where the stopping time s

$$s = \frac{\log(\beta)}{\log(1-\pi)}$$

is deduced

Overview

- Background
- Idea and Scope of the project
- Preliminary Results
- Situation in Denmark
- Non-Perfect Diagnostic Testing
- Incorporating Heterogeneity from Different Surveillance Streams

Preliminary Results

project will focus on

power function:

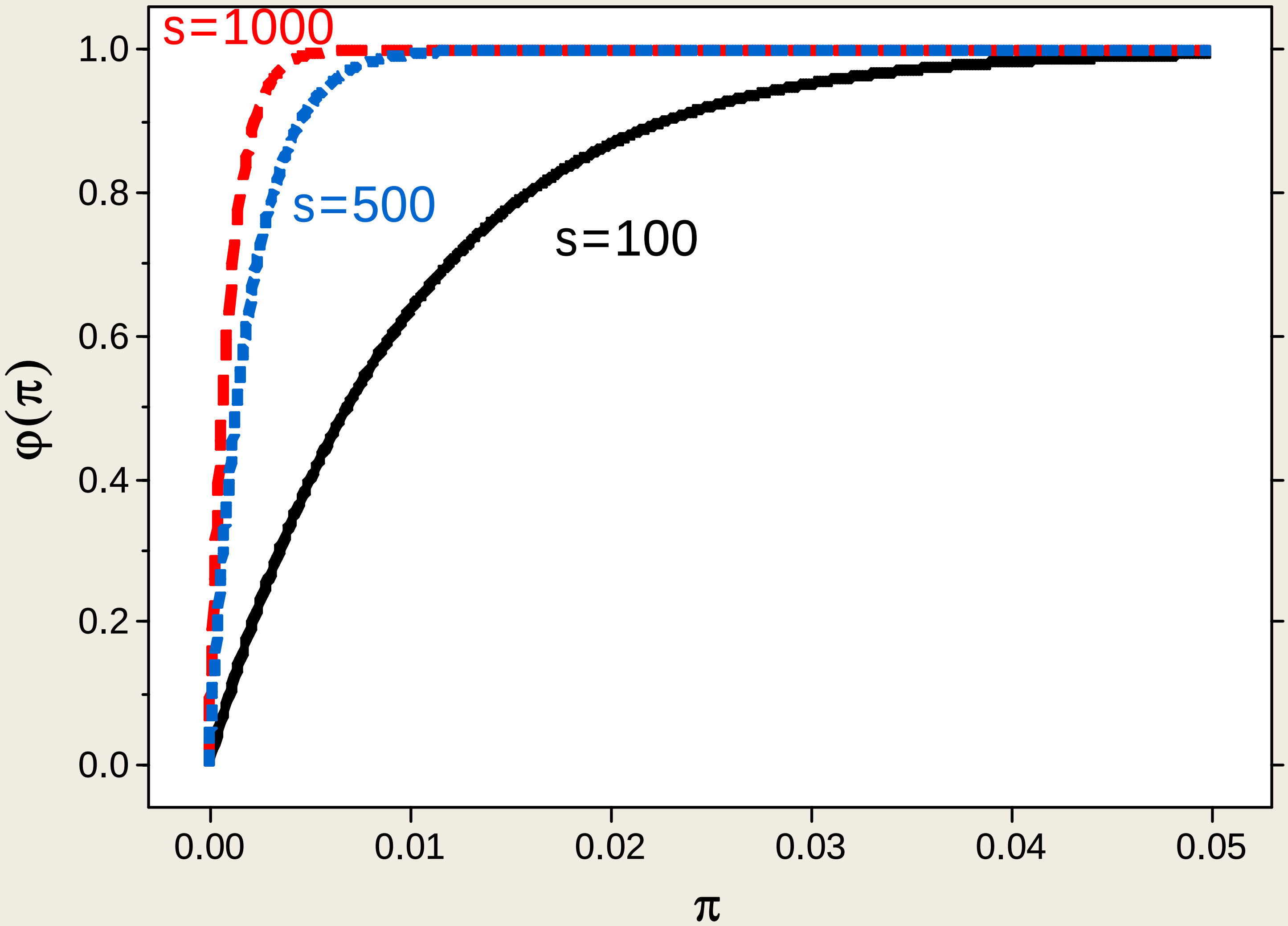
$$\varphi(\pi) = 1 - (1 - \pi)^s$$

Result: power function is monotone increasing

$$\pi_1 \leq \pi_2$$

$$\Rightarrow \varphi(\pi_1) \leq \varphi(\pi_2)$$

Monotonicity of power function



Important consequence

since true prevalence π is unknown, only minimum detectable prevalence (design prevalence) π_0 needs to be specified: it follows

$$\varphi(\pi_0) \leq \varphi(\pi)$$

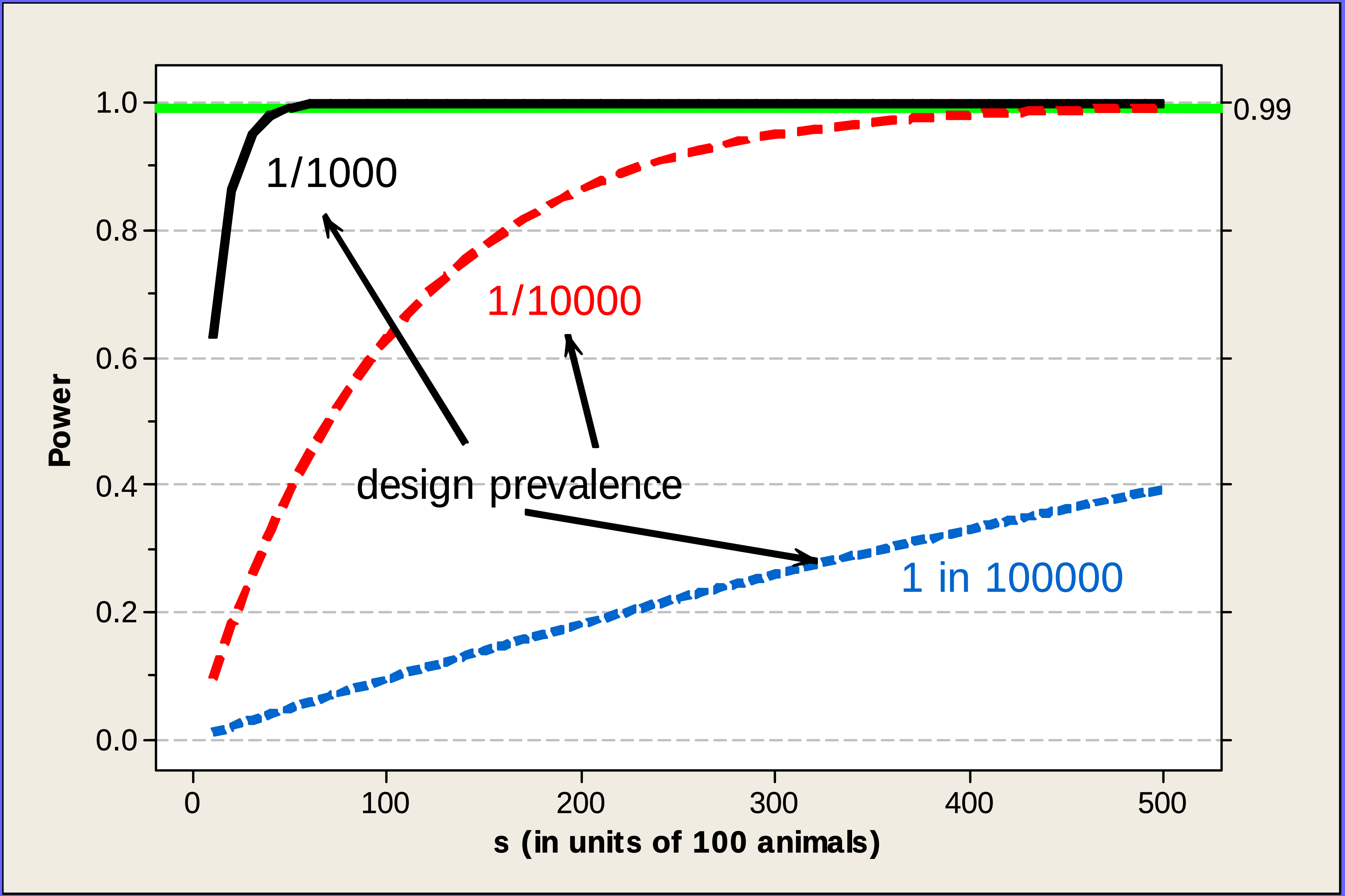
Power is also monotone in the waiting time s

power function

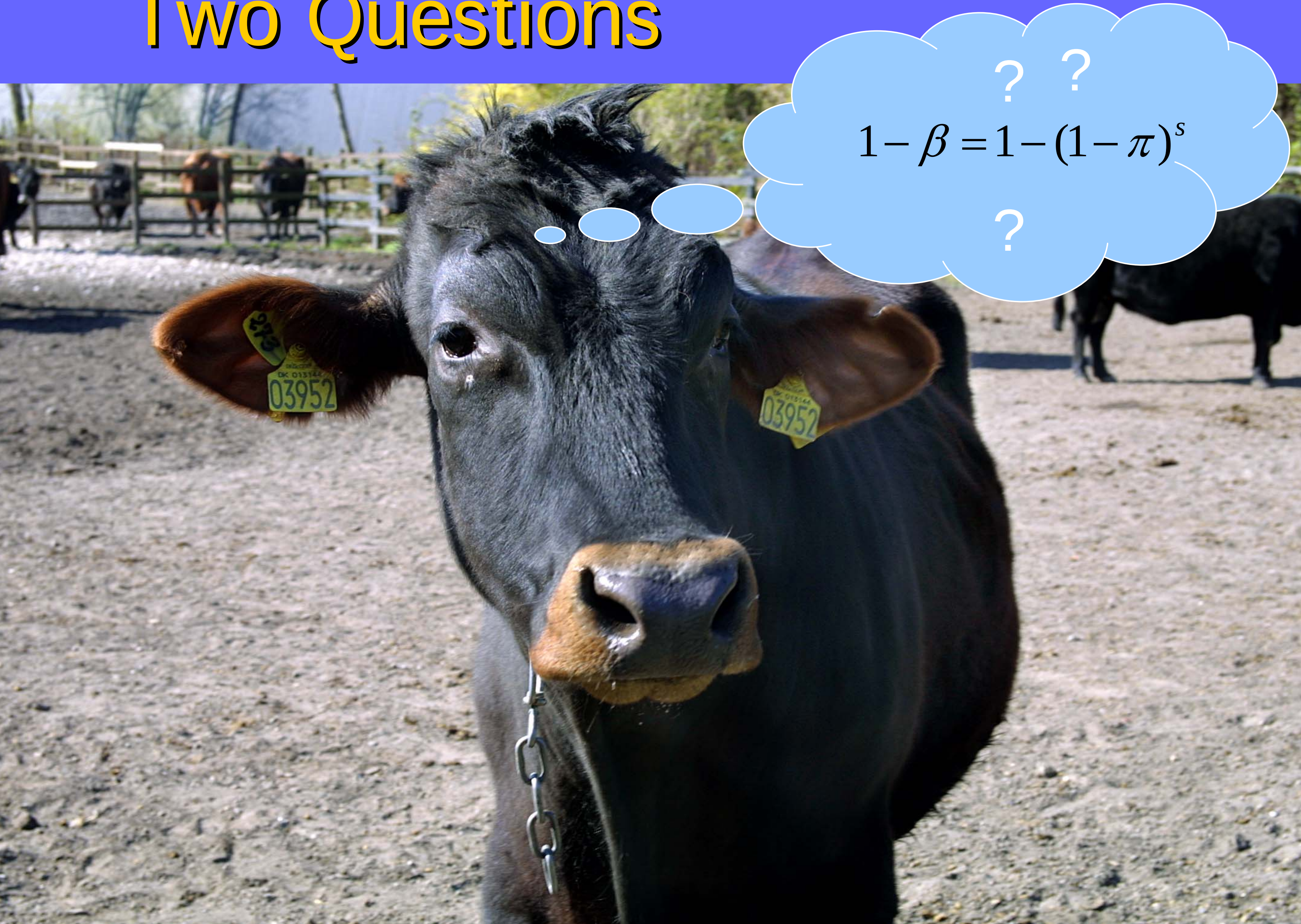
$$\varphi(s) = 1 - (1 - \pi)^s$$

(now as function of s)

Power as function of waiting time



Two Questions



What is the waiting time s to reach power of ...

$$1-\beta = 1-(1-\pi)^s \quad ?$$

from where the stopping time solution

$$s = \frac{\log(\beta)}{\log(1-\pi)}$$

is found

What is the waiting time *s* to reach power of ...

Design prevalence: 1 in	Power=0.99	Power=0.999
1000	4603	6904
10000*	46049	69074
100000*	460515	690772

* EC: Opinion in requirements for BSE/TSE Surveys, 2001

Which power have we reached given waiting time *s*?

$$\text{power} =$$

$$\varphi(\pi) = 1 - (1 - \pi)^s$$

What power is reached given waiting time *s* ?

Design prevalence: 1 in	*s*=10000	*s*=100000
1000	0.999955	1.00000
10000*	0.632139	0.99995
100000*	0.095163	0.63212

* EC: Opinion in requirements for BSE/TSE Surveys, 2001

Overview

- Background
- Idea and Scope of the project
- Preliminary Results
- Situation in Denmark
- Non-Perfect Diagnostic Testing
- Incorporating Heterogeneity from Different Surveillance Streams

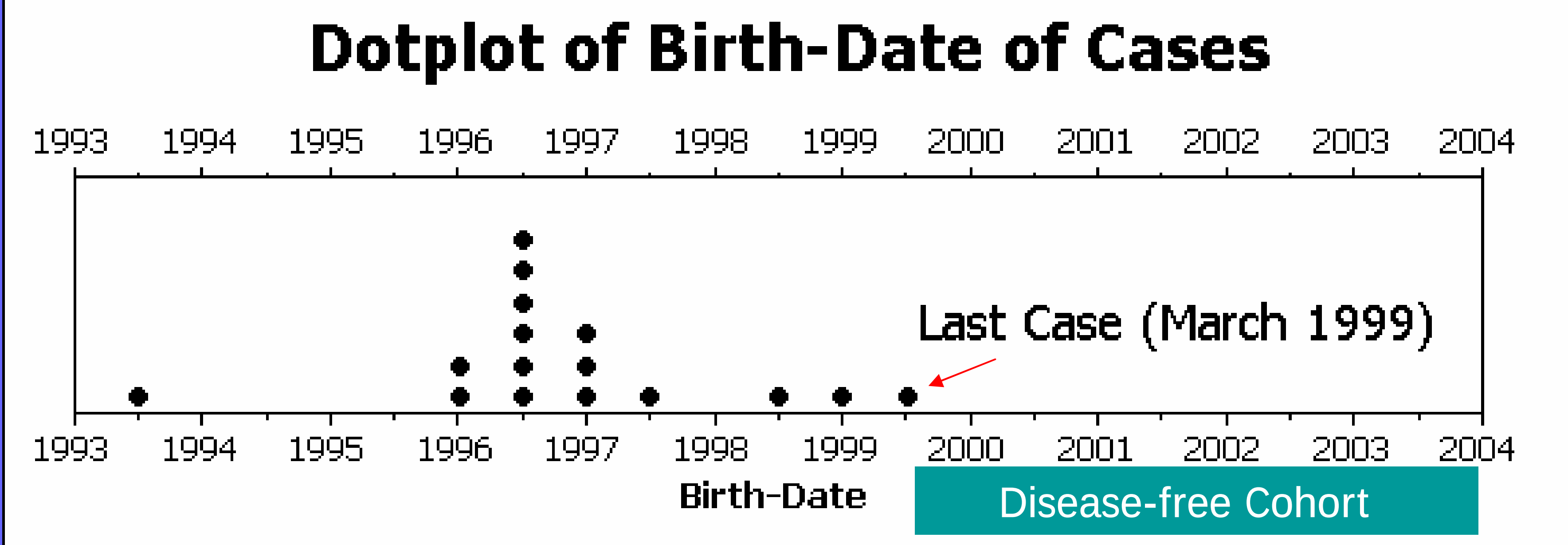
Situation in Denmark

- TSE Database: public register for BSE-testing
  - Controlled by the Danish Veterinary and Food Administration
  - Development, service and maintenance done by private company
  - Information on all animals tested for BSE since 01-Jan-2001

Situation in Denmark

- From TSE Database the following variables were made available for project:
  - Animal Identification Number
  - Age (at death)
  - Birth- and death-date
  - Cause of submission like clinical suspect, emergency slaughter, healthy slaughter,...
  - Result of BSE-testing (+/-)

Situation in Denmark: Identification of Positive Cases



Situation in Denmark

Rows: BIRTHMONTH Columns: BIRTHYEAR

	1999	2000	2001	2002	All
1	0	11154	5936	988	18078
2	0	11235	5636	692	17563
3	0	13852	6808	356	21016
4	17012	11285	6016	152	34465
5	14821	9766	4744	76	29407
6	12748	8292	3745	21	24806
7	14380	9131	3732	11	27254
8	14285	9078	3167	3	26533
9	13397	8342	2646	0	24385
10	12441	8112	2212	0	22765
11	11660	7236	1791	0	20687
12	11654	6781	1348	0	19783
All	122398	114264	47781	2299	286742

Situation in Denmark: achieved power given waiting time *s=286742*

Prevalence 1 in	Power	Prevalence 1 in	Power
10000*	1.0000	60000	0.9916
20000	1.0000	70000	0.9834
30000	0.9999	80000	0.9722
40000	0.9992	90000	0.9587
50000	0.9968	100000	0.9432

* EC: Opinion in requirements for BSE/TSE Surveys, 2001

Overview

- Background
- Idea and Scope of the project
- Preliminary Results
- Situation in Denmark
- Non-Perfect Diagnostic Testing
- Incorporating Heterogeneity from Different Surveillance Streams

π

Non-perfect diagnostic testing

Test positive/negative is not equivalent to presence/absence of disease:

$$\pi_+ = \Pr\text{ (Test positive)} < \pi$$

since $\pi_+ = \Pr(T+)$

$$= \Pr(T+ \mid D)\Pr(D) + \Pr(T+ \mid ND)\Pr(ND)$$

$$= \alpha\pi + (1-\delta)(1-\pi)$$

and, if every healthy cattle is correctly diagnosed

$$= \alpha\pi < \pi$$

where α is the test sensitivity

π

Non-perfect diagnostic testing

T waiting time for first animal testing positive: - as before -

$$\Pr(0 < T \leq s \mid \pi, \alpha > 0) = \sum_{t=1}^{s} (1-\pi_+)^{t-1}\pi_+$$

$$= 1-(1-\pi_+)^s = 1-(1-\alpha\pi)^s$$

also, $\Pr(T > s \mid \pi, \alpha > 0)$

$$= 1-\Pr(0 < T \leq s \mid \pi, \alpha > 0) = (1-\alpha\pi)^s$$

π

Non-perfect diagnostic testing

to be realistic: sensitivity will have to depend on age group:

α_a sensitivity for age group a

T_a waiting time for first animal testing positive in age group a

π

Non-perfect diagnostic testing

suppose the trial has at some given time frequencies $s_1, \ldots, s_A$ in age group $1, \ldots, A$

Power at this time?

Pr(there is a waiting time T_a s.t. $T_a \le s_a$)

$= 1 - \Pr(T_a > s_a$ for all age groups $a \mid \alpha_s, \pi)$

$$= 1 - \prod_{a=1}^{A} (1 - \alpha_a \pi)^{s_a}$$

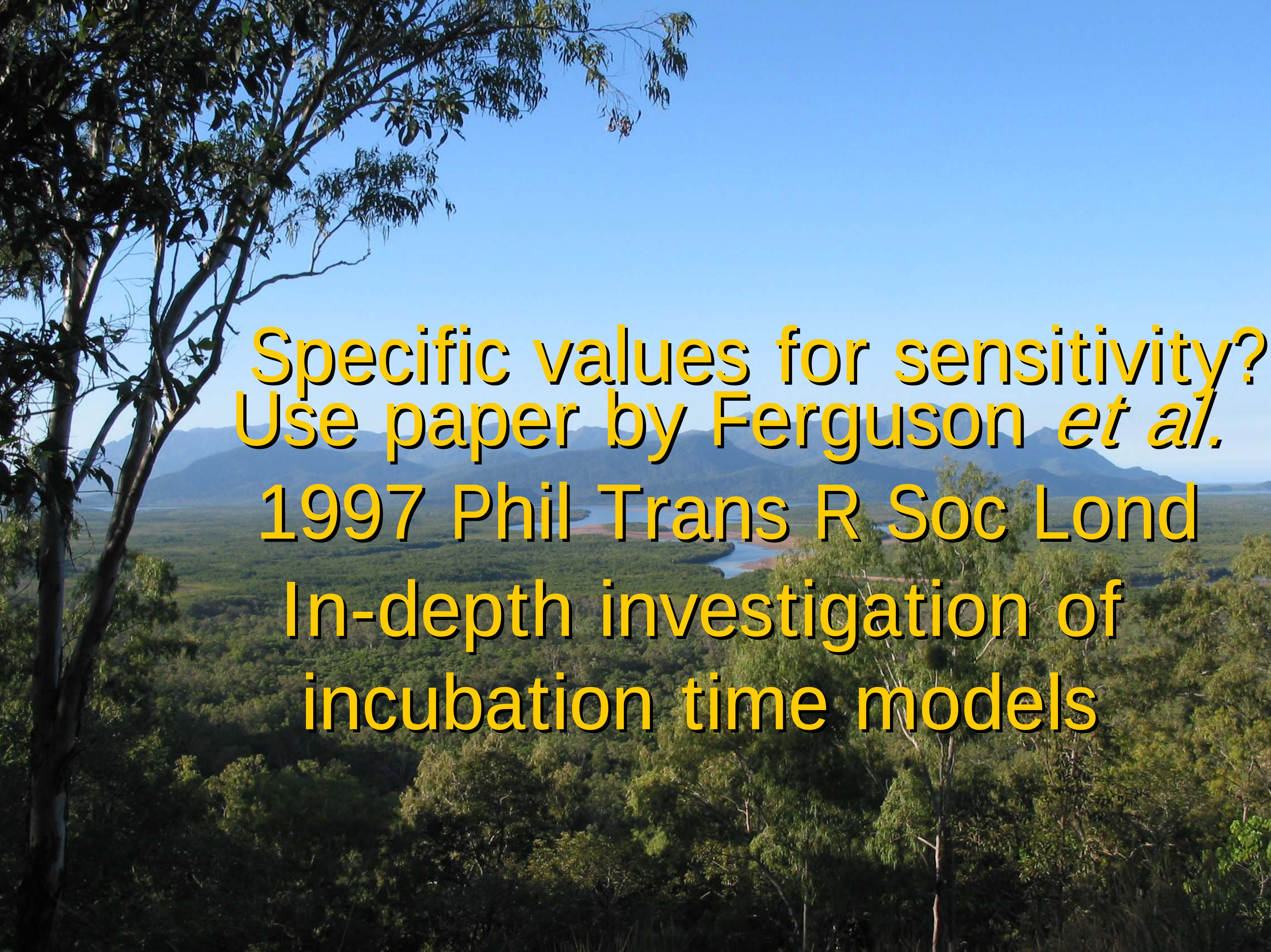
Non-perfect diagnostic testing: Situation in Denmark

Age Group	Frequency s_a	Sensitivity α_a
3	113197	0.0469
4	119439	0.2818
5	50888	0.5918
6	3218	0.8048

Situation in Denmark: achieved power incorporating sensitivity

Prevalence 1 in	Power	Prevalence 1 in	Power
10000*	0.9992	60000	0.6972
20000	0.9722	70000	0.6408
30000	0.9083	80000	0.5918
40000	0.8333	90000	0.5490
50000	0.7615	100000	0.5117

* EC: Opinion in requirements for BSE/TSE Surveys, 2001


Specific values for sensitivity?
Use paper by Ferguson *et al.*
1997 Phil Trans R Soc Lond
In-depth investigation of
incubation time models

Ferguson *et al.* 1997:

$$f(a) = \frac{1}{c}\left[\frac{\gamma_2 \exp(-a/\gamma_1)}{\gamma_3}\right]^{\gamma_2^2/\gamma_3} \times \exp\left[-\frac{\gamma_2 \exp(-a/\gamma_1)}{\gamma_3}\right]$$

where $\gamma_1, \gamma_2, \gamma_3$ are unknown parameters and c is a normalizing constant

Ferguson *et al.* 1997:

$\gamma_1, \gamma_2, \gamma_3$ are replaced by their MLEs:

$$\widehat{\gamma_1} = 1.146,\ \widehat{\gamma_2} = 0.024, \widehat{\gamma_3} = 5.71 \times 10^{-4},$$

$$\text{and } \hat{c} = 1.1350$$

Non-perfect diagnostic testing

with these values the likelihood for disease detectability within interval a to $a+1$:

$$\int_a^{a+1} f(a')da'$$

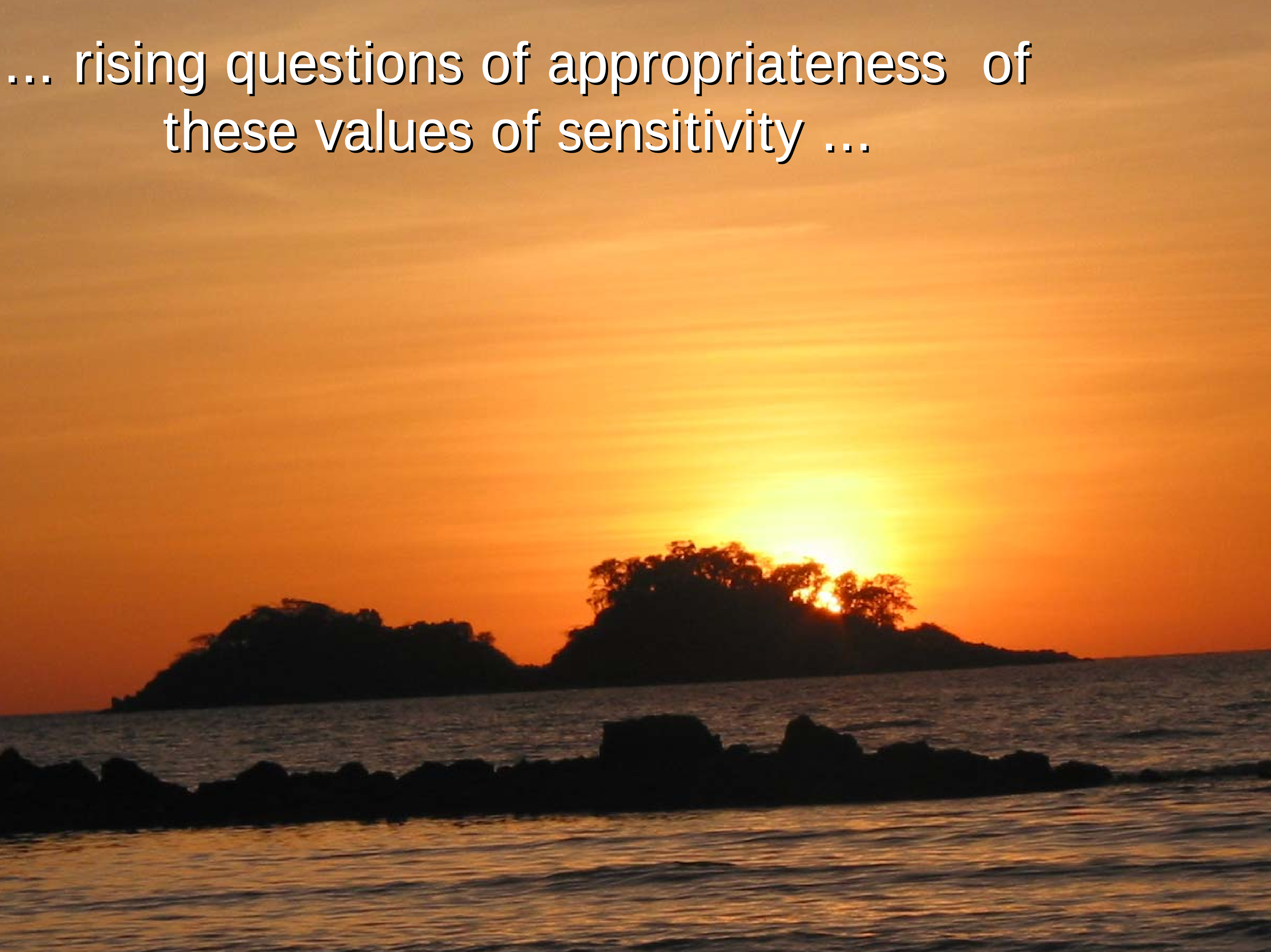
the likelihood for disease detectability up to age $\tilde{a}$:

$$\sum_{a=2}^{\tilde{a}} \int_a^{a+1} f(a')da'$$

giving

Non-perfect diagnostic testing

Age Group	$\int_{a-1}^{a} f(a')da'$	Sensitivity α_a $\sum_{a=2}^{a*} \int_{a-1}^{a} f(a')da'$
3	0.0469	0.0469
4	0.2349	0.2818
5	0.3100	0.5918
6	0.2130	0.8048


... rising questions of appropriateness of
these values of sensitivity ...

Overview

- Background
- Idea and Scope of the project
- Preliminary Results
- Situation in Denmark
- Non-Perfect Diagnostic Testing
- Incorporating Heterogeneity from Different Surveillance Streams

Incorporating heterogeneity:
important covariate: surveillance stream

Rows: Surveillance Stream Columns: Age-Years

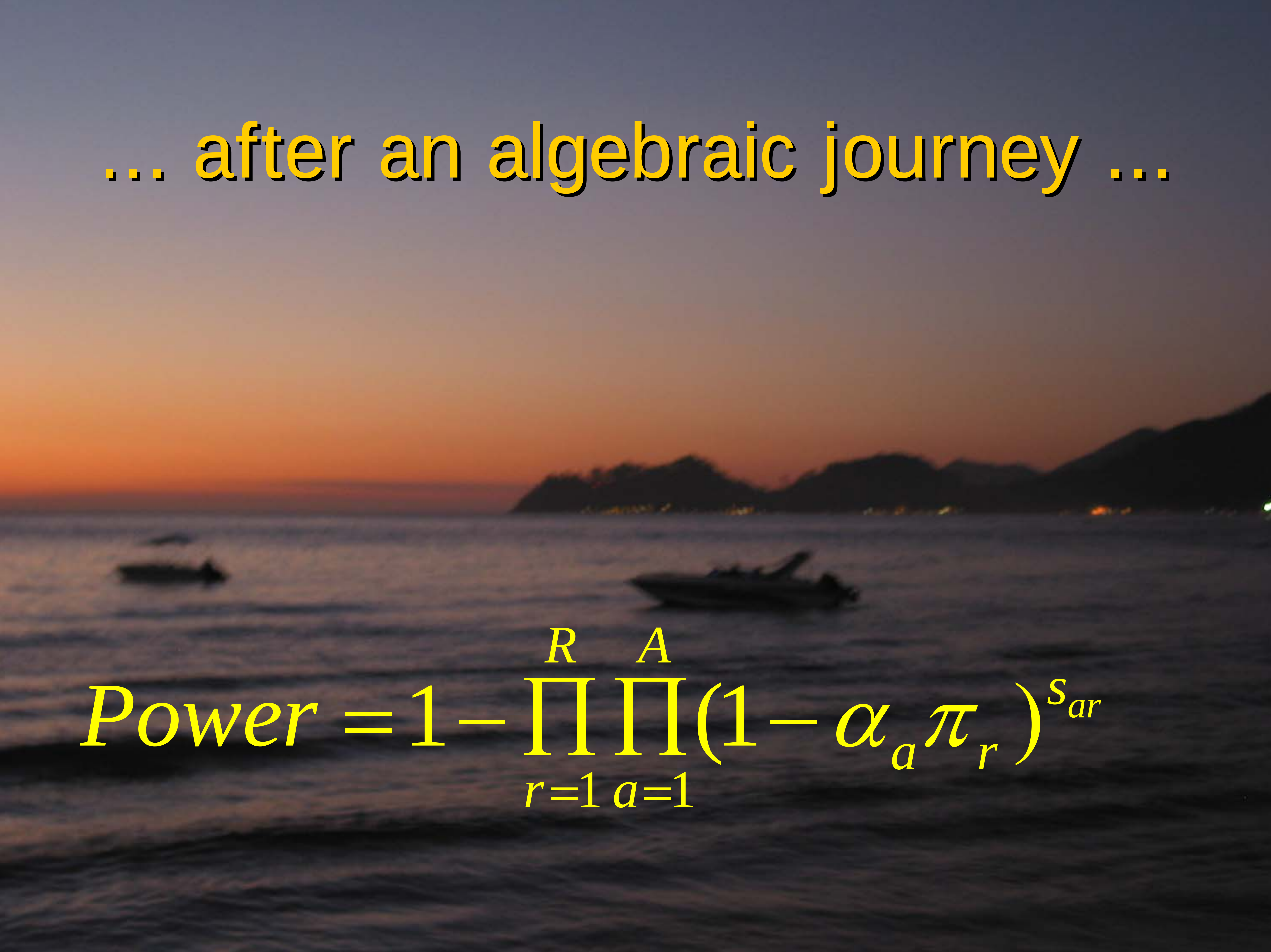
	3	4	5	6	All
HS	90511	107692	46161	3029	247393
Risk	22686	11747	4727	189	39349
All	113197	119439	50888	3218	286742

π

Incorporating heterogeneity

let s_{ar} the frequency of animals in age group a and covariate combination r,

also, let π_r denote the design prevalence in covariate combination r



… after an algebraic journey …

$$Power = 1 - \prod_{r=1}^{R} \prod_{a=1}^{A} (1 - \alpha_a \pi_r)^{s_{ar}}$$

Situation in Denmark: achieved power incorporating surveillance stream

Prevalence 1 **in HS**	Power adjusting for SS	Power not adjusting for SS
10000	1.0000	0.9992
30000	0.9991	0.9083
50000	0.9853	0.7615
80000	0.9283	0.5918
100000	0.8786	0.5117

Ratio for prevalence in risk group to healthy slaughter was estimated conservatively as 20

Discussion

- Statistical model (independence, homogeneity)
- Choice of statistical power
- Design prevalence
- Sensitivity
- …


Thank you!