



Universität Bern | Universität Zürich

**vetsuisse-fakultät**

Veterinary Public Health Institute

**Prevalence estimation from various sampling strategies in the presence of heterogeneous disease distribution between strata and imperfect diagnostic tests - results from a modelling exercise**

Marcus G. Doherr<sup>1</sup>, Nils Toft<sup>2</sup>

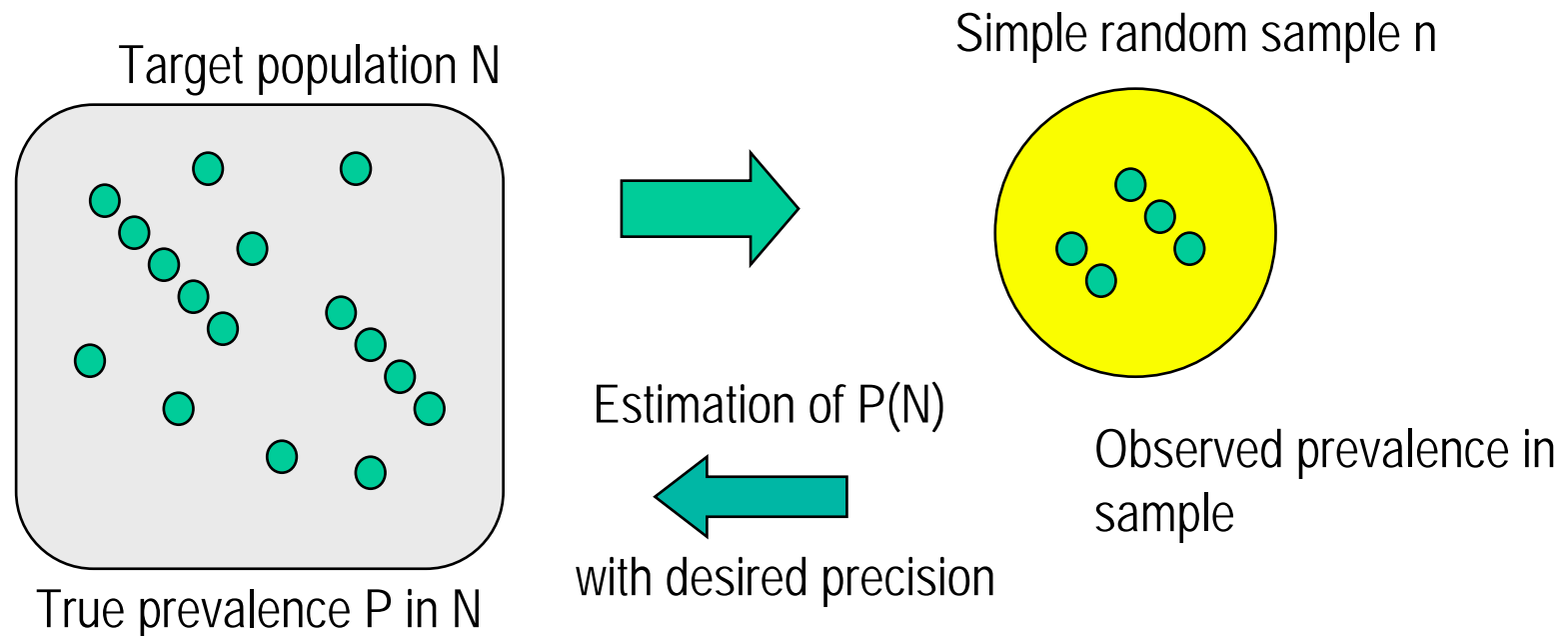
- (1) VPHI, DCR-VPH, Vetsuisse Faculty, University of Bern, CH-3001 Bern, Switzerland  
(2) Faculty of Life Science, University of Copenhagen, DK-1870 Frederiksberg C, Denmark

Surveillance Workshop  
[marcus.doherr@vphi.unibe.ch](mailto:marcus.doherr@vphi.unibe.ch)

# Content

- Brief intro
  - stratified and clustered data structures and related conventional and Bayesian prevalence estimation approaches
- Description of theoretical population and sample generating tool (R code)
- Example
  - population / sample generated
  - conventional stratum-level approach
  - two Bayesian methods ((a) 2 tests, 2 pop, no gold standard (b) 1 test, x populations, no gold standard)
- Extensions and further possibilities

# Simple (Cross sectional) Random Sampling Setup



- Sample size calculation
- **Best (MLE) estimator for prevalence**

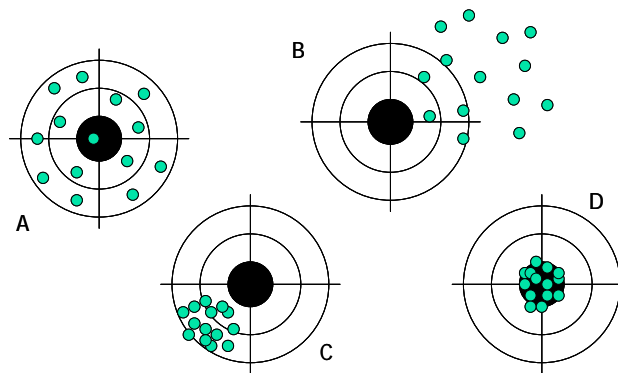
## Different sampling approaches

- Simple random sampling (SRS)
- Systematic sampling with a random seed (SYS)
- Stratified (random) sampling (STRAT)
- Cluster sampling (CLUS)
- Multiple-stage sample designs (CLUS2)
  - Cluster sampling as first step
  - Within cluster random selection using predefined criteria as a second step
- Approaches differ in their method to account for the underlying population structure, and thus the validity & precision in the prevalence estimate
  - > SAMPLE SIZE, COST / EFFICIENCY

# Sources of error in the prevalence estimation

Representativeness of „tested“ population

Measurement Error  
(Misclassification)



Random error

Analytic error

„Temporal“ error  
(lag)

# How to account for the diagnostic test system error?

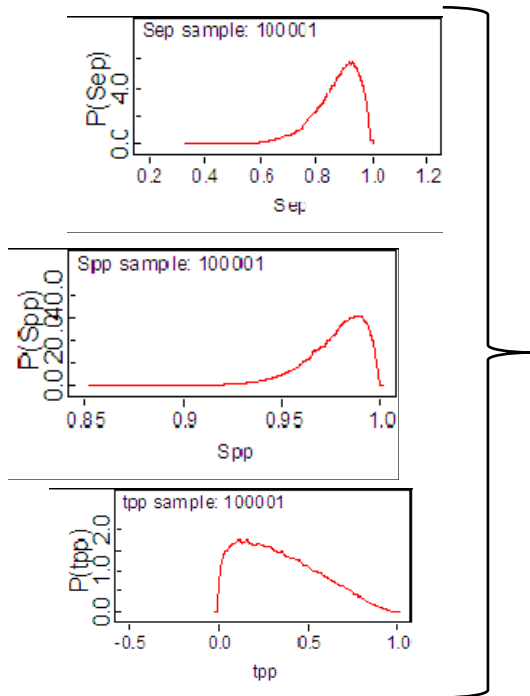
- Imperfect diagnostic tests induce a misclassification bias
  - Difference between test (apparent) and true prevalence
  - Rogan Gladen estimator

$$\text{True prevalence} = \frac{\text{AP} + \text{Specificity} - 1}{\text{Sensitivity} + \text{Specificity} - 1}$$

- Problems when prevalence is small (< 5%)
  - Test characteristics (point estimates) need to be known
- In recent years, Bayesian methods were developed to address these issues

# Bayesian prevalence estimation in WinBugs (example)

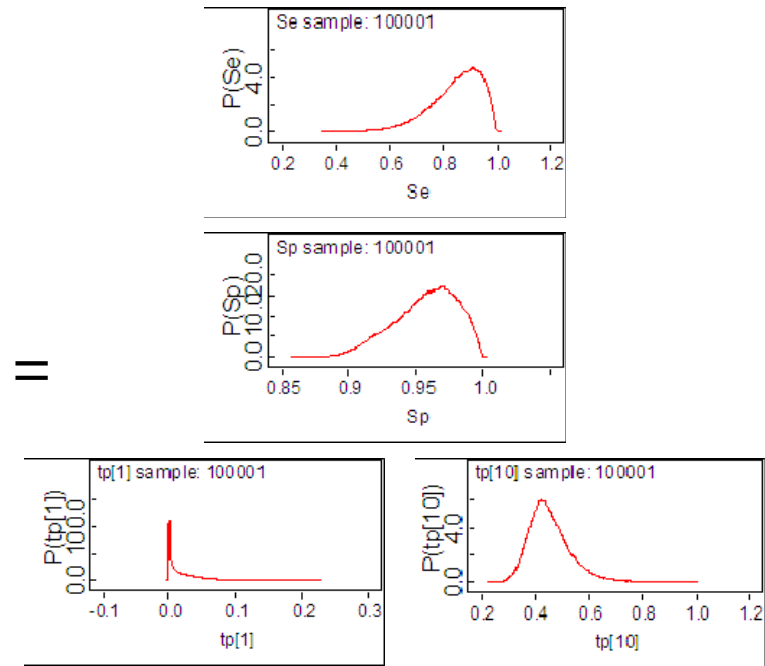
## Prior distributions



## Data and model

Flocks tested	Positive flocks
113	5
31	2
29	3
161	17
32	4
30	4
98	14
73	13
35	11
173	71
775	144

## Posterior distributions



```

model
{
  for (i in 1:k)
  {
    y[i] ~ dbin(ap[i], n[i])
    ap[i] <- tp[i]*Se + (1-tp[i])*(1-Sp)
    tp[i] ~ dbeta(alpha,beta)
  }
}
    
```

Realized in OpenBugs 2.2

110'000 iterations, the first 10'000 for the „burn in“ phase

# Validity of conventional (MLE) and Bayesian approaches

- Methods have been published and are increasingly applied on field data, but ...
  - (Parallel) gold standard test systems often not available in field situations
  - In-depth evaluation of their performance in various situations are therefore still scarce (specifically for the Bayesian methods)
- Proposal: Use of artificial data sets
  - Predefined data (population) structure with animals, herds, strata and preset prevalences at various levels
  - Predefined diagnostic test characteristics at the lowest (animal) level
  - Random (simple or structured) samples drawn from such populations

# Population & sample generator in R environment (1)

- Code developed (written) by Nils Toft (DK) in the context of a workshop in Spring 2007
- Generates herd-level and animal level artificial populations with
  - Several strata (such as regions) with different
    - herd frequencies
    - herd size distributions
    - herd-level prevalences
  - Animal level with
    - Different within-herd prevalences by animals groups (such as age)
    - True disease status (pos / neg) based on within-herd prevalence
    - Test status based on SE and SP of three diagn. tests

# 900 animals – stratified sample proportional to stratum size

## Data file structure

	region	herdid	animalid	age	truestatus	teststat1	teststat2	teststat3
1	A	102746	102746001	1	0	0	0	0
2	A	102746	102746002	1	0	0	0	0
3	A	102746	102746003	2	0	0	0	0
4	A	102746	102746004	3	0	0	0	0

Other sampling designs available

...

Etc.

## Population versus sample

Region	Animal Population			Animal sample		MLE	
	Total (N)	Pos (d)	True P	Total (n)	Pos (d)	E(P)	95%CI
A	201431	6875	0.0341	129	5	0.0388	0.128 - 0.881
B	352115	12294	0.0349	225	4	0.0178	0.005 - 0.055
C	176828	12822	0.0725	129	14	0.1085	0.061 - 0.176
D	409922	7255	0.0177	257	7	0.0272	0.011 - 0.055
E	233895	0	0.0000	161	0	0	0 - 0.023
Total	1374191	39246	0.0286	901	30	0.0333	0.023 - 0.047

# Some results from Bayesian models

			Baysian estimates based on sample					
Region	True prevalence		C/D	C/E	A-E/T1	A-E/T2	A-E/T1	A-E/T2
Stratum	Pop	Sample	E(P)	E(P)	E1(P)	E1(P)	E2(P)	E2(P)
A	0.034	0.039			0.024	0.059	<b>0.004</b>	<b>0.000</b>
B	0.035	0.018			0.020	0.049	<b>0.004</b>	<b>0.000</b>
C	0.073	0.109	<b>0.048</b>	0.122	<b>0.178</b>	<b>0.177</b>	0.109	<b>0.026</b>
D	0.018	0.027	0.035		0.052	<b>0.083</b>	0.026	<b>0.000</b>
E	0.000	0.000		(0)	0.019	<b>0.032</b>	0.003	<b>0.000</b>
Total	0.029	0.033						
		<b>Mode / GE</b>						
Test 1	SE	<b>0.600</b> / 0.400	<b>0.494</b>	0.566	<b>0.430</b>		0.561	
	SP	<b>0.995</b> / 0.950	0.995	0.992	0.996		0.992	
Test 2	SE	<b>0.995</b> / 0.950	<b>0.630</b>	<b>0.693</b>		0.984		<b>0.938</b>
	SP	<b>0.900</b> / 0.800	0.897	0.891		0.914		0.853

## Results from another exploration for a 2T2POGS Bayesian model ...

- Accuracy of estimation is highly dependent on test characteristics
  - The lower the error rates of at least one of the tests, the lesser biased the Bayesian model estimates
- Bias increases quite consistently with increasing test covariance
  - Parameters are usually overestimated
- Absolute bias of prevalence estimates is smaller for higher true prevalences
  - Accuracy increases in higher prevalence scenarios. Underestimation occurs
  - Relative bias of prevalence estimates substantially lower in higher true prevalence scenarios

## Discussion & Conclusion

- Several methods recently developed and published to estimate prevalences and test characteristics from “structured” samples
- Comparison with “truth” from an artificial data set resulted in some rather disturbing biases in the point estimates
  - Different in magnitude and direction depending on parameters and models
  - Confidence limits not yet assessed
- Further systematic work is required in order to describe in detail the advantages and disadvantages of such methods