Epidémiol. santé anim., 1997, 31-32

MULTIPLE CHANGE-POINT ANALYSIS APPLIED TO THE MONITORING OF SALMONELLA PREVALENCE IN PIGS AT SLAUGHTER AND IN THE END-PRODUCT PORK

Christensen J.[1], Rudemo M.[2]

Tous les troupeaux de porcs danois où plus de 100 porcs doivent être abattus chaque année sont suivis sur le plan sérologique grâce à des échantillons de jus de viande prélevés sur les porcs à l'abattoir pour un test ELISA (données PIG). Le contrôle de qualité de la viande de porc dans les abattoirs danois comprend un suivi permanent de la contamination du produit final - la viande de porc - par Salmonella enterica (données PORK).
L'objectif était d'envisager les variations de la prévalence des échantillons de jus de viande séro-positifs et de la prévalence de la viande de porc contaminée en fonction du calendrier, notamment pour déterminer quand les variations se produisent et pour estimer les prévalences moyennes avant et après ces variations.
La méthode du point de variation (test statistique Anderson-Darling (AD) avec l'approximation de Ornstein-Uhlenbeck du test AD) a été utilisée pour déterminer les points de variation qui ont servi à la division de la durée de l'étude en périodes où la prévalence reste constante et pour estimer les prévalences moyennes avant et après les points de variation.
Avec les données PIG, 12 points de variation ont été identifiés et les prévalences estimées entre les points de variation étaient basses pendant l'été. Les prévalences estimées à partir des données PORK étaient plus basses que les prévalences des données PIG. Les prévalences moyennes estimées allaient respectivement de 1,14% à 2,27% et de 3,51% à 7,13%. Si beaucoup de points de variation ont été identifiés - plus de quatre - comme avec les données PIG, il est apparu pratique, pour des raisons informatiques, de procéder par analyses de sous-ensembles de données. Nous avons conclu qu'avec ces données, la méthode - consistant à diviser la durée d'étude en sous-ensembles - était suffisamment solide pour donner des points de variation stables. Si les données sont très peu clairsemées, les points de variation peuvent être intégrés et conduire à la détection d'autres points de variation significatifs au plan biologique.

INTRODUCTION

The main objective of the nation-wide salmonella control program is to reduce the risk of human exposure of *Salmonella enterica* contaminated pork. The philosophy of the program is that a reduction in the salmonella occurrence in all links of the chain from stable to table is necessary for the reduction of the contamination of pork and thereby a decreased health risk for humans. Not only is it important to monitor the contamination of pork with *Salmonella enterica* but also to monitor the salmonella occurrence in the swine herds.
All Danish swine herds with an expected kill above 100 pigs per year are monitored serologically using meat juice samples from pigs at slaughter for the mix-ELISA (Mousing et al. 1997). The quality control of pork at Danish slaughter plants includes an ongoing monitoring of the contamination of the end-product — pork — with *Salmonella enterica*. The monitoring is based on culturing of samples for *Salmonella enterica* in an ongoing sampling program where a specified number of samples must be taken every 14 days at each slaughter plant. The sampled materials are: pigs (grovparteringer); cuts (finudskæringer); by-products (indmad); tongues (tunger); and processed pork (småkød og produktionskød).
There are indications that the prevalence is not constant over time. For example, the sero-prevalence was relatively low in the summer of 1995 (Carstensen and Christensen, 1996). The objective was to explore the changes in sero-prevalence of *Salmonella enterica* in meat juice samples from pigs at slaughter and in the prevalence of contaminated pork over calendar time, in particular, to detect changes and estimate the average prevalences before and after the changes.

MATERIALS AND METHODS

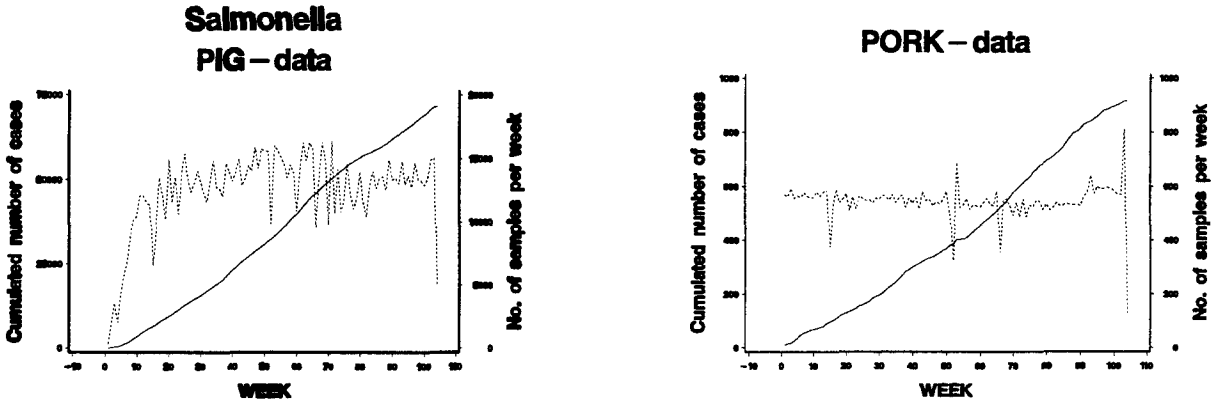
Two sources of data were used in this work: (1) the data on the serological test results of meat juice samples from pigs at slaughter originated from the Official Zoonoses Register (PIG-data); and (2) the data on the contamination of pork originated from The Danish Veterinary Service via the Database of the Danish Zoonosis Centre (PORK-data). Generally, no sampling took place on weekends, therefore the week of sampling was chosen as the unit of time. In both data sets, the number of positive samples, the number of negative samples, and the prevalence of positive samples (positive/(negative+positive)) per week was calculated and the identification variables were: key for the week, year, and the number of the week. The PIG- and PORK-data were restricted to the period from week 1 of 1995

[1] Danish Veterinary Laboratory, Bülowsvej 27, DK-1790 Copenhagen V, Denmark
[2] Royal Veterinary and Agricultural University, Department of Mathematics and Physics, Thorvaldsensvej 40, DK-1871 Copenhagen, Denmark

to week 52 of 1996. The change-point analysis (Anderson-Darling test statistic (AD) with the Ornstein-Uhlenbeck approximation of the AD) was applied to detect change-points that divided the study period into periods where the prevalence was constant and to estimate the average prevalences in these periods (Christensen and Rudemo, 1996).

The number of samples in the PIG-data was high (approximately 14,000 per week). Hence, the power of the statistical tests was expected to be high and the significance level $\alpha = 0.001$ was applied. In the PORK-data, the number of samples was much lower (approximately 550 per week) and the significance level was set to $\alpha=0.05$ or alternatively to $\alpha=0.20$.

Figure 1. The monitoring of the occurrence of *Salmonella enterica* in pigs at slaughter (PIG-data) and the end-product — pork — (PORK-data) in Denmark, January 1995 to December 1996. Left diagram: the cumulated number of sero-positive meat juice samples (cases) and the number of samples per week. Right diagram: the cumulated number of culture positive pork samples (cases) and the number of samples per week



The selection method for choosing change-points was a modified forward selection method. Briefly, the significance of previously detected change-points neighbouring the new change-point was tested given the new change-point. When more than four change-points were detected the data were divided into subsets according to change-points already detected (CP) and separate change-point analyses were performed on each subset.

As a test of the stability of the change-points CP, a subset overlapping the CP was tested for change-points. In the special case, where marginally significant change-points (p-values between 0.05 and 0.20) were added the change-points in the final model were tested individually by removing one at a time.

RESULTS

In the PIG-data, the number of samples per week ranged from 318 (week 1) to 16,397 (week 71). The number of samples per week was below 10,000 in the following periods: from week 1 to week 8 (implementation of the control program); around Easter 1995 and 1996 (week 15, 16, and 66); Christmas 1995 and 1996 (week 52 and 104); and "Bededag" 1996 (week 70) (Fig. 1). Similarily for the PORK-data, the number of samples per week ranged from 120 (week 104) to 809 (week 103). The number of samples was below 400 in the weeks 15, 52, 66, and 104. These weeks were Easter 1995, Christmas 1995, Easter 1996, and Christmas 1996, respectively (Fig. 1).

Table I. The detected change-points and prevalences in intervals between change-points. The Anderson-Darling test statistic (AD) with the Ornstein-Uhlenbeck approximation of the AD was applied

PIG -data		PORK-data			
Significance level $\alpha=0.001$		Significance level $\alpha=0.05$		Significance level $\alpha=0.20$	
Interval	*Prevalence*	*Interval*	*Prevalence*	*Interval*	*Prevalence*
1 ≤week≤ 17	5.90 %	1 ≤week≤ 56	1.33 %	1 ≤week≤ 31	1.16 %
18 ≤week≤ 36	4.85 %				
37≤week≤ 44	6.23 %			32 ≤week≤ 53	1.63 %
45≤week≤ 49	4.92 %				
50≤week≤ 54	5.57 %				
55≤week≤ 57	6.32 %			54 ≤week≤ 56	0.55 %
58≤week≤ 67	7.13 %	57 ≤week≤ 91	2.27 %	57 ≤week≤ 91	2.27 %
68≤week≤ 74	6.34 %				
75≤week≤ 77	5.53 %				
78 ≤week≤ 81	4.74 %				
82 ≤week≤ 86	3.51 %				
87≤week≤ 88	4.48 %				
89≤week≤ 104	5.29 %	92 ≤week≤ 104	1.14 %	92 ≤week≤ 104	1.14 %

In the PIG-data, 12 change-points were found (Tab. I) and the estimated prevalences between change-points were

low in summer (Fig. 2). When the first four change points were detected the data were divided into three subsets by the change-points at week 54 and 81. The change-points already detected in the analyses on the total data set were detected again in the analyses on the subsets (except the change-point between week 76 and 77) and subsequently additional change-points were detected. The test of the stability of the change-point around week 54 was tested by applying the change-point analysis to a subset of data from week 45 to week 57. The same change-points as before but no additional change-points detected. Likewise, the stability of the change-point between 81 and 82 was tested using a subset from week 75 to 86.

The estimated prevalences in PORK-data were lower than the prevalences in the PIG-data (Fig. 3). In the PORK-data only two change-points (week 56 and 91) were detected using the α=0.05 significance level but another two change-points (week 31 and 53) were added if the criterion for adding change-points was relaxed, α=0.2 (Tab.1 and Fig. 3). The change-point between week 53 and 54 was marginally significant (p-value 19%). All four change-points that were found when using α=0.2 were removed one by one to test their significance given the three other change-points. For the change-points 31, 53, 56, and 91 the p-vaules were 0.06, 0.05, 0.0006, and 0.000001, respectively.

Figure 2. The monitoring of the occurrence of *Salmonella enterica* in pigs at slaughter (PIG-data) in Denmark, January 1995, 1996, and in this figure also the begining of 1997; the average prevalences for the years 1995 (dashed line) and 1996 (solid line); and the weekly prevalences in the first weeks of 1997 (dashed line)

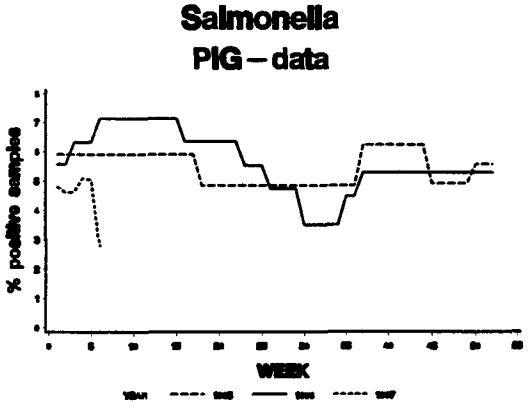
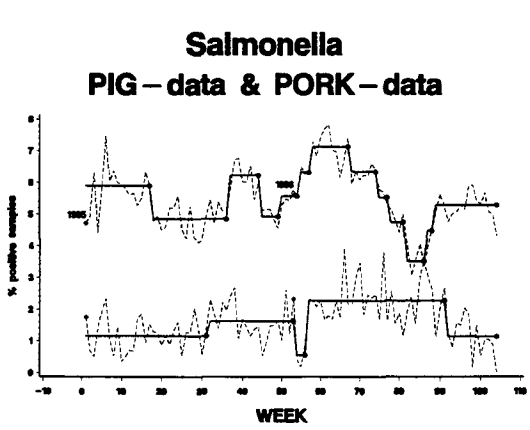


Figure 3. The monitoring of the occurrence of *Salmonella enterica* in the meat juice screening (PIG-data) and the end-product (PORK-data) in Denmark, 1995 and 1996.
Weekly prevalences of positive samples (dashed lines), the change-points and the end-points (●), the average prevalences (solid line) and the first week of the years 1995 and 1996 (◊).
The upper graph represents the PIG-data and the lower graph the PORK-data



DISCUSSION

If many —more than four— change-points were detected as in the PIG-data it was convenient for computational reasons to proceed with analyses on subsets of the data. The test of the stability of the change-point around week 54 and 81 resulted in detection of the same change-points as before but no additional change-points. We concluded that in these data the method — of dividing the study period into subsets — was robust enough to give stable change-points.

Using the significance level α=0.001 in the PIG-data resulted in detected prevalence differences of approximately 0.5%. These differences may be interpreted as biologically relevant changes when the prevalences were approximately 5%, that is a 10% change.

If the data are sparse — small weekly samples (n=500) — marginally significant change-points might be included and lead to detection of further change-points. Thus, in the PORK-data, the significance level α=0.05 resulted in the detection of a prevalence difference of 1% (a 50 % change) since the prevalences were in the range of 1% to 2%. Even smaller changes in prevalence might be biologically relevant and therefore, the significance level α=0.2 was applied to these data — with the result that changes as small as 0.5% were detected (25% change). In the actual case, the p-value for adding the change-point between week 53 and 54 was 19%, but then the weeks 54, 55 and 56 were found to constitute a period with the exceptionally low weekly prevalences 2/521=0.38%, 1/531=0.19% and 5/555=0.90%, respectively.

REFERENCES

Carstensen B., Christensen J., 1996. Fitting a random effects model for a very large dataset with censored data. Statistical modelling. Proceedings of the 11th International workshop on Statistical Moddelling, Orvieto, Italy 15-19 July 1996.

Christensen J., Rudemo M., 1996. Multiple change-point analysis of disease incidence rates. Prev. Vet. Med. 26, 53-76

Mousing J., Jensen P.T., Halgaard C., Bager F., Feld N., Nielsen B., Nielsen J.P., Bech-Nelsen S., 1997. Nation-wide *Salmonella enterica* surveillance and control in Danish slaughter swine herds. Prev. Vet. Med. In Press.