

## Opportunities and challenges when pooling milk samples using ELISA

**Græsbøll, Kaare; Andresen, Lars Ole; Hisham Beshara Halasa, Tariq; Toft, Nils**

*Published in:*  
Preventive Veterinary Medicine

*Link to article, DOI:*  
[10.1016/j.prevetmed.2016.08.001](https://doi.org/10.1016/j.prevetmed.2016.08.001)

*Publication date:*  
2017

*Document Version*  
Peer reviewed version

[Link back to DTU Orbit](#)

*Citation (APA):*

Græsbøll, K., Andresen, L. O., Hisham Beshara Halasa, T., & Toft, N. (2017). Opportunities and challenges when pooling milk samples using ELISA. *Preventive Veterinary Medicine*, 139(Part B), 93-98. DOI: 10.1016/j.prevetmed.2016.08.001

## DTU Library

Technical Information Center of Denmark

---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# 1       **Opportunities and challenges when pooling milk samples using ELISA**

2                                   Kaare Græsbøll, Lars Ole Andresen, Tariq Halasa, Nils Toft

3   National Veterinary Institute, Technical University of Denmark, Bülowsvej 27, 1870 Frederiksberg C.

4                                   [kagr@dtu.dk](mailto:kagr@dtu.dk)\*, [laoa@vet.dtu.dk](mailto:laoa@vet.dtu.dk), [tahbh@vet.dtu.dk](mailto:tahbh@vet.dtu.dk), [ntoft@vet.dtu.dk](mailto:ntoft@vet.dtu.dk)

5                                   \*corresponding author.

## 6

## 7   ABSTRACT

8   Testing large quantities of samples in order to detect one or more test-positive sample(s) is expensive  
9   and time-consuming. It is possible to optimize this process by pooling samples. Two frameworks to  
10   produce different hierarchical and non-hierarchical pooling schemes were tested and compared to  
11   standard pooling. Their efficiency and the potential savings were determined as a function of  
12   prevalence and the number of pooled samples.

13   The potential benefit of pooling samples is dependent upon the changes in the analytical sensitivity  
14   and specificity of the test used when diluting test-positive samples by pooling. To illustrate this, the  
15   sensitivity of antibody ELISA on pooled samples of bovine milk for *Salmonella Dublin*,  
16   *Mycobacterium avium spp. paratuberculosis*, and bovine virus diarrhea was tested. For these milk  
17   assays, the analytical sensitivity decreased rapidly with increasing pool sizes.

18   The efficiency of pooling is usually only measured by the number of tests performed, yet real savings  
19   depend on all the costs involved in the pooling process. These may differ between laboratories  
20   depending on the available equipment and the salaries of the technicians, among other factors.  
21   Therefore, several cost parameters were introduced to describe the total cost and thereby calculate the  
22   total savings. In terms of overall savings, both tested schemes were potentially optimal depending on  
23   the prevalence, possible pool size, and the cost of retesting. For the pool sizes of interest in this study,

24 the three-stage hierarchical pooling scheme was often marginally more efficient in terms of the total  
25 number of tests. However, if the price of re-pooling was high, the two-stage scheme performed better  
26 in terms of total savings. In addition, for low prevalences and the possibility of pooling a large  
27 number of samples, the two-stage non-hierarchical test may be more efficient, both in terms of  
28 number of tests and overall cost. In order to apply these results in different laboratory settings, a free  
29 Shiny WebApp was developed, to compare several pooling schemes with different cost parameters.

30

### 31 KEYWORDS

32 Pooling; group testing; hierarchical; shifted transversal design

33

### 34 ABBREVIATIONS

35 DD: Double Dorfman; STD: Shifted Transversal Design

36

### 37 INTRODUCTION

38 The pooling of sample material is used extensively within the veterinary field to determine farm status  
39 and/or as an indicator for further investigation. The bulk tank milk of dairy cows can be tested for  
40 pathogens such as *Salmonella Dublin* (Nielsen et al., 2005), *Streptococcus agalactiae* (Andersen et  
41 al., 2003), bovine viral diarrhoea virus (Bitsch et al., 2000), and bovine herpesvirus (Nylin et al.,  
42 2000). However, pooling has rarely been used to detect disease in individual animals.

43 The minimum number of pooled tests required to detect a single test-positive sample firstly depends  
44 upon the degree to which a test-positive sample can be diluted and still be detected: how the  
45 sensitivity changes with the pool size. Secondly, it depends upon the number of test-negative samples  
46 amongst which the test-positive sample hides: the prevalence.

47 When determining whether pools are test positive or negative, it can be beneficial to establish  
48 alternative cut-offs for the ELISA, that are lower than the defaults recommended by the manufacturer.  
49 A lower cut-off is required because pooling dilutes test-positive samples, causing a lower signal,  
50 which then decreases the sensitivity of the test. Lowering the cut-off is also a method of increasing the  
51 potential pool size.

52 Hierarchical pooling is done in multiple stages, but each sample is pooled only once per stage. For  
53 example, traditional pooling occurs in two stages, where samples are first pooled, and then samples  
54 from test-positive pools are tested individually. In a three-stage method, samples are initially pooled  
55 and samples from test-positive pools are then divided into a number of smaller pools. It is only  
56 following this second pooling stage that individual samples from positive second-stage pools are  
57 tested. This type of three-stage pooling is currently being used in HIV screening in several  
58 laboratories in the USA (Sherlock, 2007).

59 Pooling of samples can also be done in non-hierarchical structures of different dimensionality. Non-  
60 hierarchical means that all pools are tested simultaneously; ideally, it would be unnecessary to retest  
61 in order to locate test-positive samples. Examples of non-hierarchical pooling schemes are: pooling in  
62 two dimensions (which involves arranging samples in a matrix and pooling on the edges, so that each  
63 sample occurs in a row and column pool simultaneously), and arranging samples three dimensionally  
64 in a cube and then pooling on the edges where each sample occurs in three pools. Theoretically, this  
65 principle can be extended into infinitely high dimensions. However, there is a risk that the number of  
66 test-positive samples will exceed the non-hierarchical capacity of the test, and suspected test-positive  
67 samples in test-positive pools will then need to be retested individually to confirm the suspicion. An  
68 additional option of non-hierarchical schemes is to add extra pools, or 'layers' – this is known as  
69 combinatorial testing or solving the 'group-testing problem'. The advantage of using combinatorial  
70 pooling is that the probability of needing to retest samples to identify the individual test-positive  
71 sample may be significantly reduced. Reducing the number of retests can potentially save time and  
72 money.

73 The efficiency of pooling in terms of number of tests has been the subject of a number of publications,  
74 for example many of these pooling schemes are presented in Du & Hwang (2006), Hughes-Oliver  
75 (2006), and Cheng & Du (2008). The relevance of a particular pooling framework may not be solely  
76 determined by the number of tests required to detect a test-positive sample, but should also include the  
77 cost of pooling, testing, re-pooling, retesting, as well as the time taken for each of these steps.  
78 Therefore, the objective of this paper is to compare pooling frameworks in terms of all the costs in  
79 addition to the number of tests performed, in order to determine the possible opportunities in terms of  
80 savings and profits, and the challenges that must be overcome to achieve those benefits.

81

## 82 MATERIALS AND METHODS

### 83 ELISA

84 The sensitivity of commercially available ELISA on pooled bovine milk samples for the detection of  
85 antibodies for *Salmonella Dublin* (SD) (PrioCHECK Salmonella AB bovine Dublin, Prionics,  
86 Switzerland), *Mycobacterium avium* spp. *paratuberculosis* (PTB) (ID Screen Paratuberculosis  
87 Indirect, IDVet, Grabels, France), and bovine virus diarrhea (BVD) virus (Svanovir BVDV-Ab,  
88 Svanova, Uppsala, Sweden) was evaluated using the following approach. The commercially available  
89 ELISA tests were performed according to the manufacturer's instructions. Milk samples that were  
90 previously analyzed by ELISA and shown to be test positive or test negative were kindly supplied by  
91 Eurofins Steins Laboratory, Vejen, Denmark. The milk samples were taken from Danish farms as part  
92 of the surveillance programs for SD, PTB, and BVD. The numbers of test-positive milk samples  
93 included in the study were: 9 for SD, 8 for PTB and 10 for BVD. Test-positive milk samples were  
94 pooled with known test-negative milk samples, resulting in one test-positive sample being pooled  
95 with 4, 9, 24, 49, 99, 149 and 199 test-negative samples. An equal volume from each sample was used  
96 for pooling. The optical density (OD) was measured at 450nm. Results were calculated as percent  
97 positivity (PP) using Eq. (1).

$$98 \quad PP = 100 \cdot (OD_{\text{sample}} - OD_{\text{negative control}}) / (OD_{\text{positive control}} - OD_{\text{negative control}}) \quad (1)$$

99 The specificity of the ELISA and alternative cut-offs for milk samples were estimated by testing 460  
 100 known test-negative milk samples in each of the three ELISA. The alternative cut-off was calculated  
 101 as the mean percent positivity relative to the positive control of the assay, plus three times the  
 102 standard deviation.

### 103 POOLING SCHEMES

104 When presenting and deriving the pooling schemes, perfect sensitivity and specificity of the test are  
 105 assumed. The effect of pooling on sensitivity and specificity is elaborated in the discussion section.

106 1D: This is the traditional pooling scheme, and it has been used since 1915 (Hughes-Oliver, 2006).  
 107 Each sample is pooled once with a number of other samples. If a pool is test positive, all samples  
 108 belonging to this pool must be retested to identify the test-positive sample(s). In the 1D pooling  
 109 scheme, retesting is always carried out when there are test-positive pools. This means that a 1D  
 110 pooling scheme can be expensive if retesting carries a large cost. Dorfman (1943) derived an  
 111 analytical expression for the expected mean number of tests,  $E(T)$ , given the prevalence,  $p$ .

$$E_{1D}(T) = \frac{N}{n} + p' \frac{N}{n} n$$

112 Where  $N$  samples are mixed in pools of size  $n$ , the  $N/n$  pools are tested, and the samples from the  
 113  $p'N/n$  test-positive pools are tested individually (where  $p' = 1 - (1-p)^n$  is the probability of a test-  
 114 positive pool). This expression can be used to find the optimal pool size and hence the minimum  
 115 number of tests required for a given prevalence. However, this paper includes all combinations of  
 116 pool sizes and prevalences, as the optimum pool size may change according to the different costs  
 117 associated with the pooling and testing process.

118 STD: The Shifted Transversal Design was introduced by Thierry-Mieg (2006). This framework was  
 119 originally intended as a non-hierarchical scheme, meaning that test-positive samples could be found  
 120 using only one stage without retesting. However, ensuring one-stage functionality requires additional

121 pools. In this paper, the STD is tested in both a one- and two-stage mode, because the optimal mode  
122 depends on the cost of retesting. The STD includes traditional 1D testing, methods similar to the two-  
123 and three-dimensional methods mentioned in the introduction, as well as higher dimensions.  
124 Furthermore, it is a combinatorial pooling scheme, i.e. additional pools are created with a minimal co-  
125 occurrence of samples in each pool. Minimizing co-occurrence means that the STD can detect  
126 multiple test-positive samples with a reduced probability of retesting. In the STD, the number of  
127 different pooling combinations with a similar pool size increases with pool size. For a pool size of 36,  
128 there are 166 different pooling schemes of different dimensions and combinatorial complexity. The  
129 optimal pooling scheme can be determined from these by applying the costs associated with the  
130 pooling scheme steps, and comparing these costs to the individual testing. For the complete  
131 mathematical description of the STD, see Thierry-Mieg (2006). In the original paper, the STD was  
132 used in a strictly non-hierarchical (one-stage) mode by imposing the theoretical minimum number of  
133 layers given by Corollary 3 in Thierry-Mieg (2006). In this paper, the restriction is lifted and we test  
134 the STD for all number of layers from one to the maximum possible, as defined by the model.

135 When working in one-stage mode, no retest is necessary for the STD, and the efficiency is easily  
136 calculated. However, an increasing number of test-positive samples necessitate the use of the two-  
137 stage mode, and the multiple possible combinations of test-positive sample locations within the  
138 pooling structures give rise to a very complicated probabilistic structure determining the number of  
139 retests required. Therefore, the pooling schemes were simulated to determine the average number of  
140 retests. The simulation method was as follows: all possible pooling schemes with a pool size  $\leq 36$  were  
141 tested in combination with prevalences from 0.1% to 90%. For each combination of pool size and  
142 prevalence, the number of individual positive samples was drawn from a binomial distribution. This  
143 was repeated 100 times for the STD. The number of times to retest and the average number to be  
144 retested were saved for each combination of pool size and prevalence.

145 DD: This is a three-stage hierarchical pooling scheme with variable pool size. This framework will be  
146 referred to here as the ‘Double Dorfman’ (DD), as the analytical expression is based on derivations by

147 Dorfman (1943). In the first stage,  $N$  samples are mixed in pools of size  $n_1$ , the  $N/n_1$  pools are tested,  
 148 and the samples from the  $p'N/n_1$  test-positive pools are mixed again in pools of size  $n_2$ , where  
 149  $p' = 1 - (1-p)^{n_1}$  is the probability of a test-positive pool, and  $p$  is the prevalence of test positives in the  
 150  $N$  samples. Lastly, samples from the test-positive pools in the second stage are individually tested.  
 151 The expected mean number of tests,  $E_{DD}(T)$ , is expressed as:

$$E_{DD}(T) = \frac{N}{n_1} + p' \frac{N}{n_1} \left( \frac{n_1}{n_2} + n_2 \frac{n_1}{n_2} \left( 1 - \left( 1 - \frac{p}{p'} \right)^{n_2} \right) \right)$$

152 Reducing this expression and dividing by  $N$  gives the fraction of tests required,  $C$ :

$$C = \frac{1}{n_1} + p' \left( \frac{1}{n_2} + \left( 1 - \left( 1 - \frac{p}{p'} \right)^{n_2} \right) \right)$$

153 This expression only takes into account the number of tests. However, there may be further costs  
 154 associated with the additional stage of testing. Therefore, the number of pools, re-pools, and retests  
 155 were calculated for every combination of first and second pool size from 2 to 36, in combination with  
 156 prevalences from 0.1% to 90%. For an initial pool size of 36, there were 35 different stage-two pool  
 157 sizes. The optimal pooling scheme can be selected from these by applying the costs and comparing  
 158 them to the individual testing strategy.

159 Comparison of pooling schemes: When comparing pooling schemes to individual sample testing, the  
 160 simple comparison is to count the total number of tests required to detect the test-positive samples.  
 161 However, there are costs associated with the pooling itself and/or the storage and preparation involved  
 162 in retesting the individual samples identified as possible test positives by the pooling schemes. In  
 163 order to account for the price of complex pooling schemes, the costs examined were: the cost of  
 164 testing a single sample (which was set to index one); the price of retrieving a batch of samples that  
 165 were in a test-positive pool (the retrieval cost); an alternative measure of the re-pooling cost, where  
 166 the cost is per retested sample, and the hourly cost of the robot used to pool the samples. To complete  
 167 the total cost, it was also necessary to specify the time taken to pool and test the samples. When this is



168 included, it is also possible to calculate the profit gains per time unit, since a scheme that reduces the  
169 number of tests may be slower or faster in time, affecting the overall profitability. Examples of  
170 savings and profit calculations have been included in the supplementary material.

171 For simplicity, it was assumed that the test used in the simulations had 100% sensitivity and  
172 specificity (Se/Sp). The impact of non-perfect tests is elaborated in the discussion.

173 All calculations and simulations were performed using R: A Language and Environment for Statistical  
174 Computing ver. 3.1.1 (R development Core Team, 2014) in RStudio ver. 0.99.447 (RStudio Team,  
175 2015).

176

## 177 RESULTS

178 Table 1 summarizes the results of testing negative milk samples in the three ELISA. Results  
179 of pooling the nine SD test-positive milk samples with test-negative milk samples are presented in  
180 Fig. 1. Similar results were obtained when testing milk samples using assays for antibodies to PTB  
181 and BVD virus. Results show a large decrease in the percent positivity (PP) when test-positive  
182 samples are pooled with an increased number of test-negative samples. Pool sizes of more than 25  
183 samples containing one test-positive sample showed results at the level of the test-negative samples. A  
184 maximum pool size of five would give test-positive results with the alternative cut-off of 21 PP, as  
185 shown in Fig. 1. Pool sizes greater than five could result in false test-negative measurements for milk.

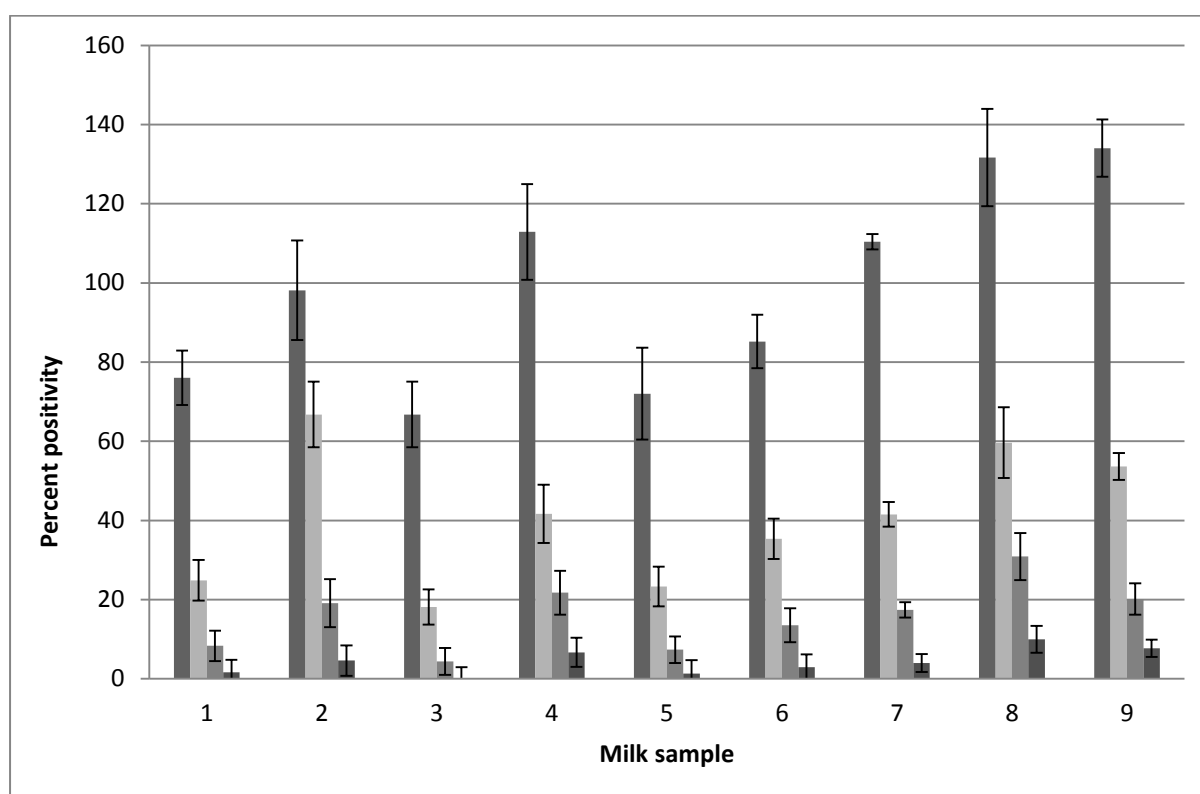
186

187 Table 1. Results from a test of the test-negative milk samples in ELISA.

ELISA for antibodies to *	Negative samples tested	Mean Percent Positivity	Standard deviation	Manufacturer's cut-off	Alternative cut-off	Specificity using alternative cut-off
<i>Salmonella Dublin</i>	460	6.65	4.62	35	21	0.99
<i>Mycobacterium avium spp. paratuberculosis</i>	460	2.00	2.18	15	9	0.995
Bovine Diarrhea Virus	460	3.50	0.93	12	7	0.99

188 \*ELISA used are presented in materials and methods.

189



190

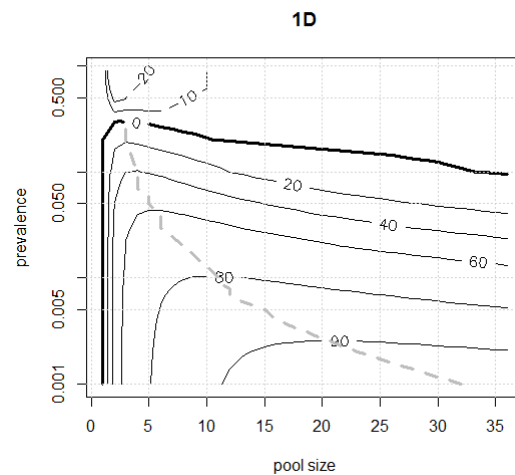
191 Fig. 1: Salmonella Dublin ELISA: Percent positivity in pooled milk samples. Nine antibody test-  
 192 positive samples and one antibody test-negative sample were pooled with known test-negative milk  
 193 samples. Results from testing the milk samples undiluted (dark gray), diluted 1:5 (light gray), 1:10  
 194 (gray) and 1:25 (black) are presented as mean values of three tests performed on separate days.  
 195 Standard errors are indicated.

196

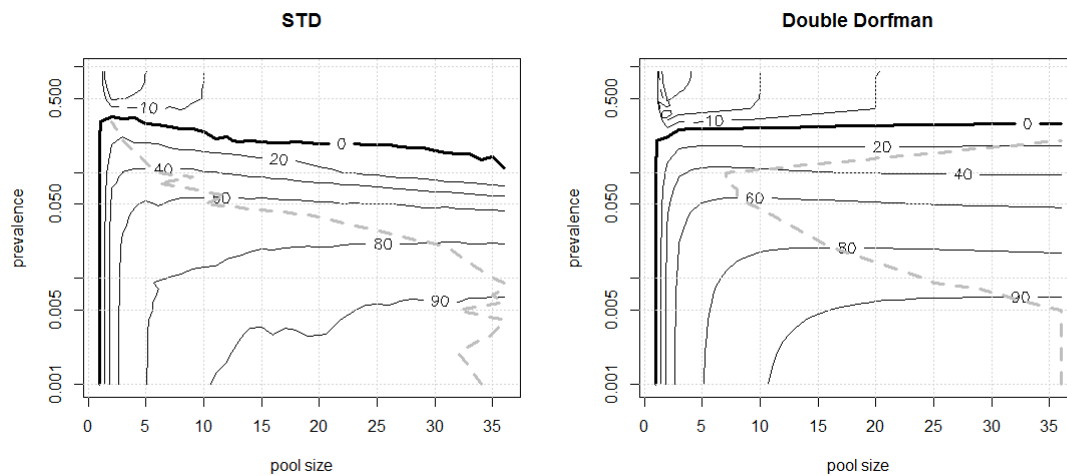
197 The results of the simulations show that the DD is more efficient when considering only the number  
198 of tests required (Figure 2). However, when introducing a retrieval cost, the STD becomes more cost-  
199 efficient than DD (Figure 3). Both the STD and DD show potentially large savings for pool sizes  
200 greater than four, providing the prevalence is low enough.

201

202

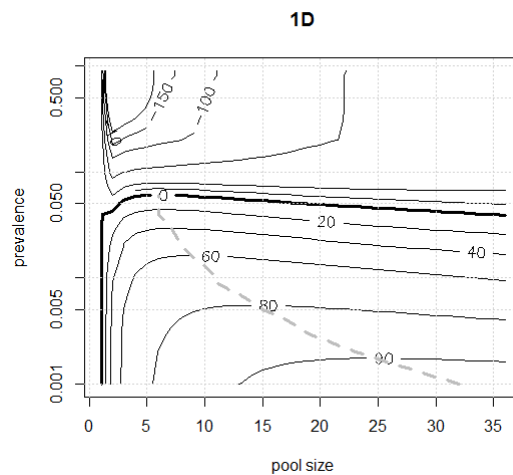


203

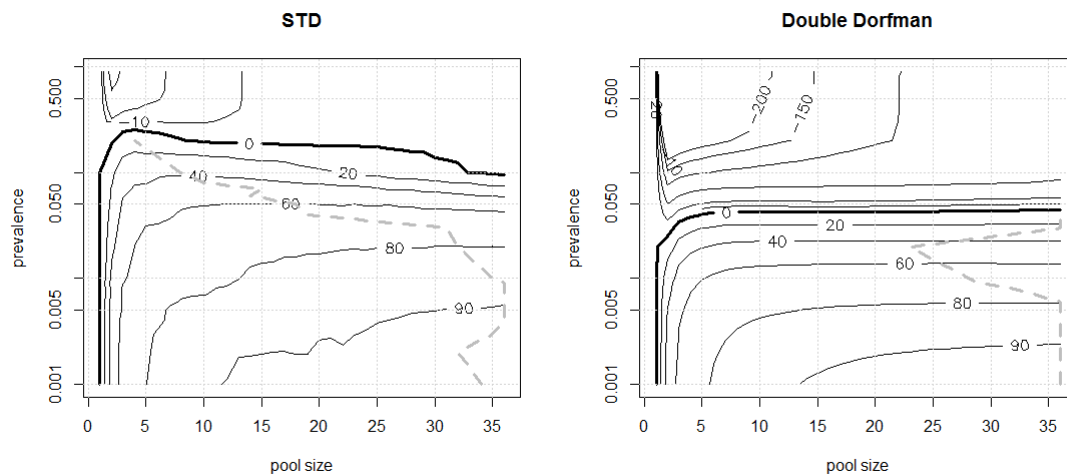


204

205 Fig. 2: Contour plots of the fraction of tests saved compared to individual testing as a function of pool  
 206 size and prevalence. The thick line indicates where the pooling scheme requires the same number of  
 207 tests as individual testing, and above this line pooling is not cost effective. The dashed gray line  
 208 indicates the optimal pool size for a given prevalence, (i.e., the pool size that gives the maximum  
 209 percentage of saved tests). This plot is equivalent to a savings plot where all costs (except the cost of  
 210 testing) are set to zero.



211



212

213 Fig. 3: Contour plots of the fraction of savings compared to the price of individual testing as a  
 214 function of pool size and prevalence, where each batch retrieval for retesting carries the cost of 10  
 215 individual tests. The thick line indicates where the pooling scheme has the same cost as individual  
 216 testing, and above this line pooling is not cost effective. The dashed gray line indicates the optimal  
 217 pool size for a given prevalence (i.e., the pool size that gives the maximum percentage of saved costs).  
 218 For all pooling schemes, there is an optimal pool size for a given prevalence (dashed gray lines in  
 219 Figures 2 and 3). It is also apparent that pooling is not efficient when the prevalence is higher than  
 220 30% (Figures 2 and 3).

221 For the small pool sizes used in this study, the DD performs best in terms of the number of tests  
222 required. The parameter space explored here covers both high prevalence and limited pool size, which  
223 includes complicated situations with multiple test-positive samples in each pool. At the other end of  
224 the parameter set, an approximate solution can be derived for the situation where the STD is better (in  
225 terms of fewer samples tested) than DD, when the pool size,  $n$ , is larger than  $((d-1)/(2p))^{2/3}$  (where  $p$  is  
226 the prevalence, and  $d$  is the dimension of the pooling scheme). This is valid when  $pn^{d/(d-1)} \ll 1$  and  
227  $n \geq 2^{d/(d-1)}$ , for any integer value of  $d \geq 2$ . Fewer samples are tested and the retrieval cost in the STD is  
228 lower, so therefore the total cost will also be lower, regardless of the specific cost parameters chosen.  
229 The mathematical derivation is available in the supplementary material.

230

## 231 DISCUSSION

232 The profitability of implementing pooling in a laboratory is dependent upon many laboratory-specific  
233 costs. In this paper, the relative cost of testing and retesting was shown to be the most important factor  
234 when determining the most profitable pooling scheme. The DD required fewer tests than the STD  
235 within the tested parameter space. However, if the costs associated with retrieval at batch level are  
236 even marginally larger than the cost of a single test, the STD scheme results in larger savings. In  
237 addition, the STD can be shown to be superior to the DD for low prevalence and large pool sizes.  
238 Both the DD and the STD are always superior to or equally efficient as the traditional 1D pooling in  
239 terms of number of tests.

240 The specific costs of testing or re-pooling/retesting may differ between laboratories. In this paper, a  
241 rather general cost structure was assumed, where the cost of testing includes the cost of a single test,  
242 the cost of a robot/technician to perform the pooling, and the time taken for both of these processes.  
243 The cost of retesting includes the cost of retrieving samples from a test-positive pool to be re-  
244 pooled/retested. The time taken to retrieve samples was not included, as it was assumed that the  
245 number of samples was large enough that samples for re-pooling/retesting could be retrieved

246 simultaneously with the testing of other samples. Therefore, the retrieval of samples is never a  
247 bottleneck, while the pooling and testing of samples was dependent upon the capacity of the facility.  
248 Although changing the parameters within the testing or retesting group changed the overall savings  
249 slightly (not shown), the only large effect was when changing the two groups relative to each other, as  
250 seen in Figures 2 and 3. It was assumed that testing samples took longer than pooling. Given that  
251 different laboratories may have different cost structures, all results from simulations were collected in  
252 a WebApp that allows for the input of cost parameters. This WebApp is available online  
253 (<https://dtuvetepi.shinyapps.io/SMARTPOOL/>) and the source code is available as supplementary  
254 material.

255 The best practice for implementing these pooling schemes may be highly dependent upon the  
256 equipment available in a specific laboratory. The DD scheme can be achieved by conventional  
257 pipetting (e.g. 'by hand'), but both the DD and the STD are more easily performed by a robot. The  
258 robot should not arrange samples in the physical structures, but rather receive a list of pools to which  
259 each sample is assigned. In this way, each individual sample is only visited once, and is directly  
260 distributed to the relevant number of pools determined by the pooling scheme. However, the best  
261 practice and the cheapest pooling method may depend on the specific laboratory.

262 From the ELISA dilution trials, it was evident that pooling impacts the Se/Sp, which were not  
263 specifically addressed in the simulations. Instead, the results of the simulations in this paper were  
264 reported as a function of the pool size, which may allow a user to choose a pool size with a desired  
265 Se/Sp based on their own dilution trials using a desired ELISA. To use the results in such a way, it  
266 must be assumed that Se/Sp is sample-specific (i.e. there is perfect repeatability): a sample that tests  
267 negative in a pooled test will test negative in all pooled tests of the same pool size, and the same for  
268 positive tests. The Se/Sp for a given disease using a specific test kit can be determined by dilution  
269 trials, as presented in this paper. Here, pooling of negatives was also used to determine an alternative  
270 cut-off in order to maximize the possible pool size. When choosing a pool size for a multi-stage  
271 pooling scheme, it can be beneficial to choose an alternative lower cut-off in the ELISA test to

272 increase sensitivity and/or pool size. The apparent loss of specificity seen when lowering the cut-off  
273 can be negated by performing the retest of individual samples using the manufacturer's cut-off. This  
274 is because false test positives will then be retested subject to the manufacturer's kit specificity. As a  
275 further means of increasing sensitivity, it may also be possible to adjust the procedure of the ELISA if  
276 there are pre-dilution steps before the OD measurements, as suggested by Brinkhof et al. (2007).

277 The dilution experiments presented in this paper are examples – a larger sample size of test-positive  
278 samples is required to give a better prediction of the change in sensitivity when pooling. Specifically,  
279 weak test-positive samples must be included in the test series in order to accurately estimate the  
280 changes in sensitivity due to pooling. Specifically, tests should be carried out to ensure that the  
281 randomness of the test does not increase around the cut-off, as reported i.e. for PCR tests with low  
282 integrity (Fleige & Pfaffl, 2006). Furthermore, it should be assessed whether samples could give rise  
283 to added unspecific reactivity when pooled.

284 Hierarchical testing was also investigated using a halving method (results not shown). This method  
285 pools samples in an initial pool, and if this initial pool tests positive, then the samples are re-pooled  
286 into two pools and tested. For each test-positive pool, the samples are divided into two new pools and  
287 tested. This is done repeatedly to the individual sample level. This method showed similar  
288 performance to the DD in terms of the number of tests, but since there are more stages (providing the  
289 initial pool size is greater than four), then this method was not superior to the DD in terms of costs.

290 Black et al. (2015) presented an optimized framework to test samples with heterogeneous prevalences  
291 at the individual level. This could be utilized within the veterinary field if risk factors were known  
292 about the individual cows. An example could be the age-dependent sensitivity of testing for  
293 *Mycobacterium avium* ssp. *paratuberculosis* (Nielsen et al., 2013). This does, however, require the  
294 extensive sorting of the samples, or the possibility of simultaneous pooling to several pools. Another  
295 use of prior knowledge may be at herd level, so that the expected prevalence in the herd (which could  
296 be determined from a previous round of testing) would determine the pooling scheme, given the  
297 results from this paper. If there is no prior knowledge of the prevalence, it could also be possible to



298 use the first-stage testing to estimate the prevalence at herd level, and use this estimate to determine  
299 the pool size of the second stage.

300 This paper expands on work presented at the annual meeting of SVEPM 2016 (Græsbøll et al., 2016),  
301 explicitly comparing 1D and higher dimensional pooling to the STD, which showed that the STD was  
302 always superior. Future work on this subject could investigate a combination of the two types of  
303 testing, including testing where prevalence is estimated after stage two to determine the pool size in  
304 stage three. It is also possible that dividing into test positive and test negative removes information  
305 from the test, and it may be advantageous and/or easier to make algorithms to identify test-positive  
306 samples based on the distribution of the continuous outcomes in a pooling scenario.

307 The code used in this paper has been integrated into a Shiny WebApp, which allows the user to  
308 specify the cost parameters and different prevalence ranges to investigate the impact on savings and  
309 profit across parameter space to emulate different laboratory settings. The WebApp is freely available  
310 at: <https://dtuvetepi.shinyapps.io/SMARTPOOL/>. The R code for the Shiny WebApp has also been  
311 included as supplementary material, so that it may be run on a local machine.

312

### 313 CONCLUSION

314 Large savings can be achieved with pooling. There is a potential to reduce the test requirements by up  
315 to 80%, even with pool sizes as small as 5, if the prevalence is low. However, there are certain  
316 restrictions that apply before those savings can be accomplished. Firstly, the expected prevalence of  
317 individual samples must be lower than 30%, otherwise pooling is not efficient. Secondly, initial tests  
318 must be performed on the specific test kit intended for use, and ideally in combination with a defined  
319 alternative cut-off to maximize  $Se/Sp$ , in order to determine how large a pool size can be used.  
320 Thirdly, samples should not display an increase in any unspecific reactivity in the test when pooled,  
321 otherwise results may be invalid. The optimal pooling framework depends on the costs associated

322 with pooling; in particular, the cost of identifying samples for retesting or re-pooling influences the  
323 optimal pooling scheme.

324

#### 325 ACKNOWLEDGEMENTS

326 All milk samples were kindly supplied by Eurofins Steins Laboratorium, Vejen, Denmark. This work  
327 was carried out as a part of the SMARTPOOL project which is funded by GUDP, grant no.  
328 34009-13-0760. The funder had no influence on the design, execution, results, or conclusions of the  
329 experiments or simulations in this study.

330

#### 331 DECLARATIONS

332 Some of the work referred to in this manuscript has previously been presented by the authors at the  
333 annual meeting of SVEPM 2016, which includes a conference paper.

334

#### 335 REFERENCES

336 Andersen, H. J., Pedersen, L. H., Aarestrup, F. M., & Chriél, M. (2003). Evaluation of the  
337 surveillance program of *Streptococcus agalactiae* in Danish dairy herds. *Journal of dairy science*,  
338 86(4), 1233-1239.

339 Barillot, E., Lacroix, B., & Cohen, D. (1991). Theoretical analysis of library screening using a N-  
340 dimensional pooling strategy. *Nucleic acids research*, 19(22), 6241-6247

341 Bitsch, V., Hansen, K. E., & Rønsholt, L. (2000). Experiences from the Danish programme for  
342 eradication of bovine virus diarrhoea (BVD) 1994–1998 with special reference to legislation and  
343 causes of infection. *Veterinary microbiology*, 77(1), 137-143.

- 344 Black, M.S., Bilder, C.R., & Tebbs, J.M. (2015). Optimal retesting configurations for hierarchical  
345 group testing. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*
- 346 Brinkhof, J.M.A., Houwers, D.J., & Van Maanen, C. (2007). Development of a sample pooling  
347 strategy for the serodiagnosis of small ruminant lentiviral infections using the ELITEST-MVV  
348 ELISA. *Small Ruminant Research*, 70(2), 194-199
- 349 Cheng, Y., & Du, D. Z. (2008). New constructions of one-and two-stage pooling designs. *Journal of*  
350 *Computational Biology*, 15(2), 195-205.
- 351 Dorfman, R. (1943) The detection of defective members of large populations. *Ann. Math. Statist.*, 14,  
352 436–440.
- 353 Du, D., & Hwang, F. (2006). Pooling designs and nonadaptive group testing: important tools for DNA  
354 sequencing (Vol. 18). World Scientific Publishing Company Incorporated.
- 355 Fleige, S., & Pfaffl, M. W. (2006). RNA integrity and the effect on the real-time qRT-PCR  
356 performance. *Molecular aspects of medicine*, 27(2), 126-139.
- 357 Græsboøll, K, Andresen, L.O., Halasa, T., Toft, N. How few pooled tests are needed to detect a single  
358 positive sample? SVEPM 2016 proceedings.
- 359 Hughes-Oliver, J. (2006). Pooling experiments for blood screening and drug discovery. *Screening:*  
360 *Methods for Experimentation in Industry, Drug Discovery, and Genetics*, Springer NY, 48-68
- 361 Nielsen, L.R., & Ersboøll, A.K. (2005). Factors associated with variation in bulk-tank-milk *Salmonella*  
362 Dublin ELISA ODC% in dairy herds. *Prev. Vet. Med.*, 68(2), 165-179
- 363 Nielsen, S. S., Toft, N., & Okura, H. (2013). Dynamics of specific anti-*Mycobacterium avium* subsp.  
364 paratuberculosis antibody response through age. *PLoS ONE*, 8, e63009

- 365 Nylín, B., Strøger, U. and Rønsholt, L. (2000). A retrospective evaluation of a Bovine Herpesvirus-1  
366 (BHV-1) antibody ELISA on bulk-tank milk samples for classification of the BHV-1 status of Danish  
367 dairy herds. *Prev. Vet. Med.* 47, 91-105
- 368 R Development Core Team. (2014). R: A Language and Environment for Statistical Computing. R  
369 Foundation for Statistical Computing. Vienna, Austria. [www.R-project.org](http://www.R-project.org)
- 370 RStudio Team. (2015). RStudio: Integrated Development Environment for R. RStudio, Inc. Boston,  
371 MA. [www.rstudio.com](http://www.rstudio.com)
- 372 Sherlock, M., Zetola, N. and Klausner, J. (2007) Routine detection of acute HIV infection through  
373 RNA pooling: Survey of current practice in the United States. *Sexually Transmitted Dis.*, 34, 314–316.
- 374 Thierry-Mieg, N. (2006). A new pooling strategy for high-throughput screening: the Shifted  
375 Transversal Design. *BMC bioinformatics*, 7(1), 28
- 376