Implementation of sample pooling to strengthen the self-monitoring in the food industry: Case study of mycotoxins in cereals

Jean-Michel Galharret, Benjamin Mahieu, jérémy ratel, Evangelia Krystalli, Katerina Pissaridi, Evelyne Vigneau, Erwan Engel

► To cite this version:

Jean-Michel Galharret, Benjamin Mahieu, jérémy ratel, Evangelia Krystalli, Katerina Pissaridi, et al.. Implementation of sample pooling to strengthen the self-monitoring in the food industry: Case study of mycotoxins in cereals. Food Research International, 2025, 10.1016/j.foodres.2025.115937 . hal-04940588

HAL Id: hal-04940588
https://hal.inrae.fr/hal-04940588v1

Submitted on 13 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Journal Pre-proofs

Implementation of sample pooling to strengthen the self-monitoring in the food industry: Case study of mycotoxins in cereals

Jean-Michel Galharret, Benjamin Mahieu, Jérémy Ratel, Evangelia Krystalli, Katerina Pissaridi, Evelyne Vigneau, Erwan Engel

PII: S0963-9969(25)00274-1
DOI: https://doi.org/10.1016/j.foodres.2025.115937
Reference: FRIN 115937

To appear in: *Food Research International*

Received Date: 4 October 2024
Revised Date: 3 February 2025
Accepted Date: 4 February 2025

Please cite this article as: Galharret, J-M., Mahieu, B., Ratel, J., Krystalli, E., Pissaridi, K., Vigneau, E., Engel, E., Implementation of sample pooling to strengthen the self-monitoring in the food industry: Case study of mycotoxins in cereals, *Food Research International* (2025), doi: https://doi.org/10.1016/j.foodres.2025.115937

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2025 Published by Elsevier Ltd.

1 # Implementation of sample pooling to strengthen the self-monitoring
2 in the food industry: Case study of mycotoxins in cereals

3 Jean-Michel Galharret[1], Benjamin Mahieu[1], Jérémy Ratel[2], Evangelia Krystalli[3], Katerina
4 Pissaridi[3], Evelyne Vigneau[1], Erwan Engel[2*]

5

6 1 Oniris, INRAE, StatSC, F-44300 Nantes, France

7 2 INRAE, UR QuaPA, MASS Group, F-63122 Saint-Genès-Champanelle, France

8 3 Yiotis Anonimos Emporiki & Viomixaniki Etaireia, Athens, Greece

9

10 *Corresponding author: Erwan Engel,

11 INRAE, QuaPA-MASS group, 63122 Saint-Genès-Champanelle, France.

12 E-mail: erwan.engel@inrae.fr

13

14 ## Abstract

15 Implementation of sample pooling strategy in the food chemical surveillance could lead to a
16 strengthening of the food safety by increasing the number of analyzable samples. The analysis
17 of a pool of samples and no longer individual samples was applied to the issue of self-
18 monitoring mycotoxins in cereal-based foods, thanks to a data set provided by the surveillance
19 based on ELISA-kits of two mycotoxins – zearalenone (1121 samples) and ochratoxin A (1601
20 samples) – in four different types of cereal products. After fitting the distribution of mycotoxin
21 concentrations determined in this product category by a Pareto distribution and considering the
22 measurement error in the decision threshold, numerical simulations of pooling were
23 implemented using the Dorfman-2-step strategy. Simulations showed promising results for
24 three out of the four case-studies of zearalenone and ochratoxin A. While being as sensitive
25 and specific as the current one-by-one system, the pooling approach led to a reduction of the
26 number of analyzes performed by 75 – 87 % in three out of the four case studies. Nevertheless,
27 in unfavorable analytical conditions, the pooling approach can lead to an increase of the total
28 number of analyzes.

29 Keywords: chemical hazard ; food safety surveillance ; zearalenone ; ochratoxin A; ; numerical
30 simulations ; corrected maximum limit ; Dorfman strategy

31 ## 1 Introduction

32 Particularly when it aims to protect sensitive populations such as children of less than three
33 years of age for whom chemical exposure during the first 1000 days of life is essentially of
34 food origin and will play a key role in health well beyond the childhood (ANSES, 2016),
35 surveillance of risky chemical contaminants must be strengthened. This involves intensifying

36 the frequency and extent of surveillance carried out by health authorities but also providing
37 food manufacturers with the means to develop or intensify self-monitoring at critical control
38 points in their processing chain. To do this, it is not only necessary to develop detection tools
39 that are fast, inexpensive and sufficiently efficient but also to propose ingenious strategies to
40 optimally organize their implementation (Engel et al., 2022).

41 Mycotoxins are chemical agricultural contaminants that are naturally produced by several
42 fungal species, mainly *Aspergillus, Penicillium and Fusarium*. Control measures involve pre-
43 and post- harvest practices, as contamination may occur upon cultivation and storage. Even if
44 mycotoxins occur in various food commodities, the main exposure sources for humans are
45 cereals such as wheat, maize, rice and sorghum and nuts (Battilani et al., 2020; Eskola et al.,
46 2020). Exposure to mycotoxins via cereals is all the more critical for infants and children as
47 the extreme importance of this food category in child nutrition (Pereira et al., 2022). According
48 to the literature on the occurrence of mycotoxins in cereals, ochratoxins and zearalenone are
49 part of the most commonly detected mycotoxins (Wan et al., 2021). Ochratoxin A (OTA) is
50 the most toxic member in ochratoxin family and classified as a potential human carcinogen
51 (group 2B) by the International Agency for Research on Cancer. Little is known about the
52 toxicological effects of chronic exposure to low-level OTA in foodstuffs, but some recent
53 studies highlight the toxicological and physiological effects of successive exposure to OTA at
54 food regulatory limits (Wei et al., 2021). Zearalenone (ZEA) is a non-steroidal estrogenic toxin,
55 with toxic hazards including reproductive toxicity, digestive toxicity, hepatotoxicity and
56 immunotoxicity (Cai et al., 2024). Considering hazard to human health, these mycotoxins are
57 part of the mycotoxins for which maximum permissible levels in food have been set in Europe
58 (EU 2023/915).

59 For the food industry, mycotoxins monitoring in the incoming raw materials is thus important
60 and unavoidable to ensure food safety, as it is practically impossible to minimize contamination
61 upon processing. Taking into consideration the serious impact of mycotoxins on human health,
62 especially in childhood, the difficulties in eliminating contamination of crops, and the high rate
63 of incidents, testing for mycotoxins, and especially ZEA and OTA, on cereal food products on
64 a regular basis is indispensable. Hence, in many cases, a strict sampling plan is applied in all
65 cereal-based raw materials, involving control for mycotoxins at every receipt and resulting in
66 a huge number of samples to be analyzed.

67 In addition to the development of increasingly rapid while remaining inexpensive tools such as
68 Enzyme-Linked Immunosorbent Assay (ELISA) kits for monitoring these mycotoxins on an
69 industrial scale, sample pooling approaches could make it possible to meet this need to monitor
70 continually a very large number of samples. Sample pooling is a group testing strategy
71 introduced by Dorfman (1943) where multiple samples are combined into a pool that is tested
72 using a single analysis (Aldridge et al., 2019). This strategy aims to exclude a large number of
73 negative samples and detect a few positive samples at lower cost (Cheng et al 2020). Sample
74 pooling has already been experimented in areas where positive samples are rare, such as high-
75 throughput testing of drugs (Hann et al., 1999), detection of pathogens in clinical samples
76 (Currie et al., 2004) or in food products (Jagadeesan et al., 2019) but it is above all the COVID
77 crisis which has highlighted the benefits of this approach for large-scale safety surveillance
78 issues (Bilder et al., 2020; Mallapaty, 2020; Verdun et al., 2021). While in terms of food safety
79 the method was mainly devoted to microbiological hazards, a first group sought to evaluate the
80 benefits of sample pooling to improve the monitoring of priority chemical hazards (Cheng et
81 al., 2020). To do this, they use existing databases to model mycotoxin contamination in single
82 corn kernels by fitting statistical distributions to experimental data and used Monte-Carlo

Journal Pre-proofs

83 simulation to evaluate assay sensitivity, specificity as well as potential gains compared to a
84 single test approach. Starting with the screening of mycotoxins by ELISA methods as a case
85 study, Cheng et al., (2020) showed that the sample pooling strategy could be interesting for
86 reducing laboratory costs (volume of solvent, pipetting). Nevertheless the advantage was
87 generally mixed, in particular due to the often high rate of samples exceeding the regulatory
88 maximum limits set for the agricultural raw material studied together with the high measure
89 uncertainty due to the heterogeneity of mycotoxin partition in whole grain samples compared
90 to more processed samples like powders (Brodal et al., 2020; Köppen et al., 2010). By focusing
91 on the case of regulatory surveillance of PCBs in pork meat for which the ML exceedance rates
92 were of the order of 0.1% that is much lower than in the former work of Cheng et al (2020),
93 Ratel et al (2024) recently demonstrated the full potential of the approach. Based on two
94 lognormal distributions for fitting the uncontaminated and contaminated data respectively, a
95 numeric simulation demonstrated that with an optimal pool size of 25, the pooling strategy
96 enabled to decrease the number of analysis per sample by 19 while preserving the sensitivity
97 and specificity of non-conformity diagnosis.

98 After pointing out the relevance of sample pooling to strengthen surveillance conducted by
99 health authorities (Ratel et al., 2024), this new study aims to assess its potential to help food
100 industry manufacturers optimize self-monitoring. To do this, it draws on unique databases of
101 ELISA kit surveillance of two mycotoxins - zearalenone (1121 samples) and ochratoxin A
102 (1601 samples) - in 4 different types of cereal products including either incoming raw materials
103 or processed finished products intended for human consumption or infant nutrition. To evaluate
104 the implementation of the pooling strategy, the distributions of mycotoxin concentrations in
105 flours are fitted from the available data by selecting a probability distribution that best fitted
106 the data with particular attention to the tails of the distribution. This model makes it possible
107 to simulate pooling of n samples by using the Dorfman-2 step strategy (Ratel et al., 2024). In
108 order to better control the sensitivity associated with the decision threshold at the level of the
109 pools, the choice of the new "maximum corrected limits" is then discussed. Finally, simulation
110 studies are performed for evaluating the performance of sample pooling strategy by considering
111 both sensitivity, specificity and, of course, its cost-effectiveness in terms of relative number of
112 tests per sample.

113

114 2 Material and methods

115 2.1 Samples

116 In total 1121 samples were analyzed for Zearalenone (ZEA) from February 2017 till September
117 2023 and 1601 samples for Ochratoxin A (OTA) from February 2017 till January 2023. These
118 samples either were incoming raw materials (521 samples) such as wheat, rye, barley and rice
119 flour, semolina, and cornstarch collected from various food industries in Greece or processed
120 final products (600 samples) such as cereal snacks, breakfast cereals and cereal-based baby
121 food randomly collected from the Greek market. For the study, samples were categorized
122 according to the Maximum Level (ML) permitted for each mycotoxin according to Regulation
123 (EU) 2023/915 on maximum levels for certain contaminants. For ZEA, 3 cases were formed
124 with ML 20, 50 and 100 µg/kg, while for OTA one case with ML 3 µg/kg.

125 2.2 Reagents and instrumentation

126 The quality of methanol reagent was from Fisher Scientific (Hampton, USA). Ultrapure water
127 (resistivity 18.2 MΩ cm-1) was produced by a Milli-Q purification system (Purite). The
128 RIDASCREEN® Ochratoxin A 30/15 enzyme immunoassay kit was from R-Biopharm AG
129 (Darmstadt, Germany), while the Bio-Shield Zearalenone kit was from Prognosis Biotech
130 (Larissa, Greece). The microplate spectrophotometer was provided by BIO-TEK (Santa Clara,
131 USA).

132
2.3 Sample analysis

133 Samples were analyzed following the protocol of the commercial kits RIDASCREEN®
134 Ochratoxin A 30/15 enzyme immunoassay kit (Giray et al., 2009) and Bio-Shield Zearalenone
135 kit (Trevisi et al., 2020). The method principle was the competitive enzyme immunoassay. In
136 brief, the wells in the microtiter strips of the kit were already coated with specific antibodies
137 against each specific mycotoxin. Mycotoxin standards or sample solutions and enzyme
138 conjugate were added. Free and enzyme conjugated mycotoxin compete for the specific
139 mycotoxin antibody binding sites. Any unbound enzyme conjugate was then removed in a
140 washing step. Enzyme substrate and chromogen were added to the wells and incubated. Bound
141 enzyme conjugate converts the colorless chromogen into a blue product. The addition of the
142 stop solution lead to a color change from blue to yellow. The measurement was made
143 photometrically, in a microplate absorbance reader (BIO-TEK, Santa Clara, USA.) at 450 nm
144 for both ZEA and OTA. The absorption was inversely proportional to the mycotoxin
145 concentration in the sample.

146
2.4 Method performance

147 The limit of quantification (LOQ) for each parameter was determined as ten times the average
148 absorbance plus the standard deviation of the signal of ten blank samples. The relative standard
149 deviation for each analyte and for each contamination level was calculated from 40 samples
150 analyzed in two different days. The results are summarized in Table 1.

151

152
3 Theory/calculation

153 In the following, the contaminant concentration is denoted by X. X_i is the value of X on an
154 individual sample i. A sample i is claimed contaminated when

155 $$X_i > ML$$

156 where ML is the maximum authorized concentration. Therefore, the prevalence of
157 contaminated sample denoted by $prev$ is the proportion of sample for which $X_i > ML$.

158
3.1 Modeling the contaminant concentration

159 To take into account the presence of extreme values of X, its distribution is often chosen from
160 heavy-tailed distributions (e.g. log-normal distribution, Pareto distribution). Moreover, when
161 the analytical method has a limit of quantification (LOQ), the concentration X is left-censored,
162 so the Pareto distribution is a good candidate (Merritt, 1898) as it has two parameters to be
163 determined, namely λ and x_m, enabling to manage both prevalence through ML and the

Journal Pre-proofs

164 probability of a measure to be lower than LOQ. x_m can be interpreted as the censoring
165 parameter and λ as the shape parameter (called the tail index).

166 These parameters are estimated based on the following conditions:

167 - The proportion of values below LOQ matches observed proportion in databases.
168 - The proportion of values above ML matches the expected prevalence of contaminated
169 data.

170 The log-normal distribution could be an alternative to the Pareto as it is also defined from two
171 parameters and it is often used to model PCBs contaminant distributions (Ratel et al., 2024).
172 However, with mycotoxins, the Pareto distribution seems more suitable. Indeed, PCBs are
173 ubiquitous in the environment and are sometimes overexpressed due to environmental
174 overexposure of animals during their breeding. Therefore, the distribution of their
175 concentrations in animal products such as meat classically fits a lognormal distribution (e.g.
176 Berninger & Tillitt, 2019). For mycotoxins, a slightly different model can be expected because
177 they are produced by fungi that may have contaminated the cereals and develop and produce
178 these toxins during storage.

179 ## 3.2 Include measurement error in the decision rule

180 If the contaminant concentration X was measured without error, the decision rule might be: the
181 sample i is declared to be contaminated if its concentration value X_i is greater than ML. Note
182 that, for error-free X values, this decision rule has a sensitivity (i.e. true positive rate,
183 probability of being detected positive when truly contaminated) and a specificity (true negative
184 rate, probability of not being detected positive when truly uncontaminated) equal to 1. In this
185 situation, the ML corresponds to the threshold for the decision rule.

186 In practice, for an individual sample i, the measurement X_i^* is given with a measurement error
187 whose value is proportional to the true value X_i. This aspect is considered in the modeling by
188 assuming that:

189 $$X_i^* = X_i + \varepsilon_i$$

190 where ε_i has a Gaussian distribution with mean 0 and standard error σ_i proportional to X_i (i.e.
191 $\sigma_i = RSD \times X_i$, where RSD stands for the Relative Standard Deviation). If the previous
192 decision is used, the sensitivity and the specificity are unknown since a contaminated value
193 near the ML can have a measured value under the ML (see Figure 1).

194 As the sensitivity is the ability of a test to detect correctly the contaminated sample, it is a key
195 parameter to be controlled. Therefore, in this work, the decision rule is defined so that the
196 corresponding sensitivity (denoted by Se) reaches a user defined value. This can be achieved
197 by setting an optimal threshold (denoted by corrected Maximum Limit, ML^*) that verifies:

198 $$P(X_i^* > ML^* | X_i > ML) = Se.$$

199 Since ML^* has no analytical expression, it is estimated by Monte Carlo simulations.

200 Figure 2 shows the evolution of the corrected Maximum Limit, ML^*, for a sensitivity target of
201 99% and ML fixed at 20 (as in the case ZEA1) with a RSD ranging from 0 to 0.20 and
202 prevalence ranging from 0.1% to 10%. As it can be observed, the corrected Maximum Limit
203 decreases more or less linearly as the measurement uncertainty increases. The lower the
204 prevalence, the more the ML* is far below ML. These observations are an illustration of the
205 level of confidence required for decision making when testing is carried out under complex
206 conditions (relatively large RSD and low prevalence of contamination).

207 **3.3 Sample pooling strategy**

208 As said in the Introduction, the detection of contaminated samples is based on the Dorfman
209 procedure (Dorfman, 1943). With N samples to test, in the first step, samples are gathered into
210 pools of size n_P. Having a pool made of n_P samples whose individual contaminant
211 concentration values are X_1,...,X_{n_P}, the contaminant concentration of the pool is their average
212 $X_P = \frac{\sum_{i=1}^{n_P} X_i}{n_P}$. The pool is submitted to the analytical instrument (here ELISA-kit). The observed
213 measurement, X_P^*, follows the same distribution as the individual measurement value i.e.

214 $$X_P^* = X_P + \varepsilon_P$$

215 where ε_P has a Gaussian distribution with mean 0 and standard deviation equal to $RSD \times X_P$.
216 The decision rule for the pool is based on a threshold value ML_P^* which is determined to reach
217 a fixed value for the pool sensitivity Se_P. This threshold ML_P^* defined such that:

218 $$P(X_P^* > ML_P^* | \text{at least one } X_i \text{ is greater than ML}) = Se_P.$$

219 Pools for which the X_P^* measurement are below ML_P^* are considered to be made up of solely
220 uncontaminated samples.

221 On the contrary, when X_P^* measurement is above ML_P^*, the pool is likely to contain at least one
222 contaminated sample. Thus, all the samples into such a pool are individually tested in the
223 second step of the Dorfman procedure. In the second step, the decision rule for each individual
224 sample is the one described in section 3.2.

225 At the end of this strategy, the global sensitivity is the product of the sensitivities of both steps,
226 i.e. $Se_P * Se$. For instance, for an overall sensitivity of 0.99 (i.e. 99%), setting the sensitivity
227 of the pool sample at 99.5% ($Se_P = Se = \sqrt{0.99} = 0.995$) balances the risk. Note that, in
228 practice, a major drawback of the Dorfman procedure is to assume that the test has a perfect
229 sensitivity (see for instance Jaszkiewicz, 2020). In the proposed procedure, this assumption is
230 not required and the sensitivity of the global procedure is a target for the optimization of the
231 corrected maximum limit.

232 **3.4 Simulations**

233 To investigate the interest of the Dorfman strategy, numerical simulations were performed. The
234 simulation process was as follows. First, N random values were drawn from a Pareto
235 distribution. These values represent the true values, X_i of samples' contaminant concentration.

Journal Pre-proofs

236 In this paper, a size of N= 50 was considered but realistic analytical batches may be smaller.
237 These N true values were then gathered in pools of size 1 up to N. Note that in case the pool
238 size is not a divisor of N, the last pool contains less samples. For each constituted pool, the
239 average of the true values from the samples that belong to the pool, X_P, was computed. A
240 Gaussian noise was added to the average value to simulate the measurement X_P^* (see section
241 3.3). This measurement was then compared to the corresponding pool threshold, ML_P^*, to take
242 the decision of whether or not the pool contains at least one contaminated sample. If the
243 decision is that the pool indeed contains at least one contaminated sample, then the same
244 process as the one described for the pool (measure and decision) is applied to the individual
245 samples pertaining to this pool. At the end of the process the number of measures performed,
246 the sensitivity and the specificity are collected. This entire process is repeated a large number
247 of times (say 10 000) with the different parameters of the case studies described in Table 1 in
248 terms of LOQ, ML, etc.

249 # 4 Results and Discussion

250 ## 4.1 Distribution fitting

251 Figure 3 provides the fitting of the Pareto and Log-Normal distributions on the data. It shows
252 that the Pareto distribution was more adapted to model mycotoxin contaminations than the
253 Log-Normal distribution in all four cases. The parameter of censor x_m ranged from 5.8 to 7.5
254 for ZEA datasets and was equal to 0.2 for OTA dataset and the parameters of shape from 2 to
255 3 for all four samples.

256 ## 4.2 Corrected Maximum Limits

257 The corrected Maximum Limits, ML_p^*, were determined for the four case studies investigated
258 (ZEA I, ZEA II, ZEA III, OTA). As shown in section 3, the Maximum Limit (ML), the
259 prevalence ($prev$) and the relative measurement uncertainty (RSD) had an impact on the
260 performance of the test, and especially in terms of sensitivity. For global sensitivity objective,
261 fixed here to 99%, corrected Maximum Limits ML_p^*, can be assessed as a function of the pool
262 size n_p, as shown in Table 2.

263 The decision rule for pools in the first step of the Dorfman procedure is based on such
264 thresholds. In the specific case of a pool size equal to 1, Table 2 gives the corrected Maximum
265 Limits to be considered in the second step of the Dorfman procedure, or if the sampling
266 protocol is defined without pooling.

267 Table 2 shows that for the ZEA I case study, the corrected Maximum Limit was below the LOQ
268 value for pools of any size. The pooling strategy was therefore unsuitable in this situation. In
269 fact, ZEA I case study presented rather specific conditions with a ML value that was not high
270 compared to the LOQ ($ML/LOQ = 2$) and a relative measurement uncertainty RSD of 35%. For
271 the OTA case study, the corrected Maximum Limit was always larger than or equal to LOQ
272 but the choice of large pool size may be questionable because the corrected Maximum Limit
273 becomes close or even equal to the LOQ. In the other case studies, the corrected Maximum
274 Limits became lower than LOQ when the pool size became too large. Obviously, the lower the
275 LOQ as compared to the ML, the larger the pool size can be.

276 In order to compare the corrected Maximum Limit proposed in Table 2 for decision making,
277 let us consider a simpler rule of thumb. The naive way of selecting limits for decision making,

Journal Pre-proofs

considering the dilution effect, is to compare the analytical result for a pool with the ML value divided by the number of samples in the pool, i.e. to define the limit as $ML_p^0 = \frac{ML}{n_p}$.

Figure 4 illustrates ML_p^* and ML_p^0 as a function of the pool size considering the parameters of the ZEA II case. Note that the larger the pool size, the smaller ML_p^0 was compared to ML_p^* except when samples were not pooled (i.e. pool size = 1). For instance, with pools of four samples, the proposed corrected Maximum Limit, ML_p^*, would be 16.7, and the naïve corrected Maximum Limit, ML_p^0, equal to 12.5. Monte Carlo simulations with pools of size 4 showed that the sensitivity using ML_p^* was 98.8%. This result was expected as ML_p^* was set to give a global sensitivity of 99%. When using ML_p^0, the estimated sensitivity was only 89.0%, which means that approximately 11% of the contaminated samples could not be detected. However, test specificity was very high at 99.9% for both limits.

4.3 Average number of tests expected with the pooling strategy

In order to benchmark the "sample pooling" approach with the classic approach consisting of analyzing all samples individually, the simulation results obtained with Dorfman pooling method were examined for the four case studies of mycotoxin monitoring using ELISA-kits. Criteria used were the sensitivity, the specificity and the savings obtained thanks to the pooling, assessed by the number of tests performed per sample (ratio of the total number of tests to the number of samples present in the pool).

Regarding the sensitivity (data not shown), when the Corrected Maximum Limit (ML_p^*) was considered as decision limit, it achieved the global sensitivity target, which was set at 99% for all pool sizes and all case studies. Regarding the specificity (data not shown), it was equal to 1 for all pool sizes, in the case studies of OTA (prevalence of 0.1%) and ZEA in products II and III (prevalence of 0.2% and 0.1%, respectively). In contrast, the specificity in the case study of ZEA in product I (prevalence of 8%) specificity was between 89% and 92%, depending on the size of the pool considered. Note that these results guarantee the internal validation of the proposed procedure. However, an interesting avenue for future research could be an external validation using spiked experimental samples with a known level of contamination.

Figure 5 shows the reduction of the number of analysis per sample expected using the Dorfman pooling procedure, when considering ML_p^* as a decision limit. According to the prevalence estimated for each mycotoxin, the optimal pool size was different depending on the type of mycotoxin and the type of cereal-based food considered. In the ZEA I case study, the use of sample pooling was not recommended since the ratios of the total number of tests to the number of samples present in the tested pools were always strictly greater than 1. In this case the use of pooling would be even less efficient than individual analysis of samples. These results can be explained by the fact that the analysis conditions did not meet two of the main criteria for the applicability of sample pooling recalled by Malapaty (2021) and more recently by Ratel et al (2024): a relatively low prevalence and a sufficient margin between the sensitivity of the analytical method used (in this study the ELISA kits) and the ML set for the contaminant considered (in this case ZEA) in the matrices studied, modulo the uncertainty of the analytical method. Indeed, as shown in Table 1, the prevalence of exceeding the ML was logically higher than for the two other case studies concerning ZEA (8% for ZEA I vs 0.2% for ZEA II and <0.1% for ZEA III) because the ML set for infant food (category I) was stricter in order to ensure the protection of a more sensitive population, infants (Engel et al., 2022). Moreover, for ZEA I case study, the margin between the ML (20 µg/kg) and the LOQ of the method (10

322 µg/kg) was all the more restricted as the uncertainty of measurement U=2×RSD=34.8% was
323 relatively large.

324 In contrast, Figure 5 shows that for ZEA II and III case studies, the relative number of tests per
325 sample was less than one, indicating that a significant gain could be expected when pooling
326 was implemented. For ZEA II, the use of pools of 5 samples required the implementation of
327 25% of tests compared to 100% with a sample analysis one by one, i.e. a saving of 3 tests out
328 of 4 (75%), if the possible additional logistical costs linked to the implementation of pooling
329 was neglected. For ZEA III, a cost saving of approximately 87% can be achieved for pools of
330 10 or 13 samples. Concretely, this means that, considering a batch of 20 samples, out of 20
331 analyzes traditionally carried out one by one, sample pooling led to only 2 or 3 analyzes being
332 carried out, on average. It is also interesting to note that the performance of sample pooling
333 remained fairly stable for pool sizes between 10 and 20. With the aim of maximizing the speed
334 of obtaining control results, and therefore speed up batch release and improve the flexibility of
335 analyzes more generally, the pooling of 20 samples could even be considered. For OTA case
336 study, the optimal pool size equal to 9 samples would make it possible to reach a relative saving
337 of 82%, i.e. barely lower than for the ZEA III case.

338 It is interesting to confirm that if the ZEA II, ZEA III and OTA case studies were much more
339 favorable to the use of sample pooling than the ZEA I case, it was in particular because they
340 fulfilled its applicability criteria much better than ZEA I. Prevalence rates, margin between ML
341 and LOQ as well as the analytical uncertainty (RSD) are criteria of main importance. It
342 therefore seems interesting to continue developments to improve the performance of analytical
343 methods, whether by lowering the LOQ or by limiting measurement uncertainties in order to
344 improve performance and the relative gains expected to sample pooling implementation. In this
345 regard, it is worth noting that the analytical cost saving could be already significantly increased
346 with the use of the most recent ELISA kits for ZEA detection, whose performance in terms of
347 LOQ is greatly improved (for example LOQ of 1.75 µg/kg for the RIDASCREEN® kits).
348 Furthermore, even if our results obtained from cereal-based products collected on the Greek
349 market can be considered robust and realistic in view of the number of samples analyzed (1121
350 and 1601 samples for ZEA and OTA, respectively) and the collection period (6 years), the
351 parameterization of the sample pooling model could be readjusted in certain specific situations
352 likely to impact mycotoxin contamination levels in cereal-based products (geographic location,
353 climate change...).

354 5- Conclusion

355 This study showed that sample pooling, which had already proven to be a serious avenue for
356 strengthening the monitoring of certain key chemical contaminants by food safety authorities
357 (Ratel et al, 2024), is also a very promising approach for improving the effectiveness of self-
358 monitoring that food companies work on contaminants. Taking as an example the monitoring
359 of two mycotoxins in cereal products by ELISA test kits, the simulations showed that for the
360 most favorable products, this approach could make it possible to reduce by 75 to 87% the
361 number of analyzes carried out by analyzing the samples.

362 This article proposes a new approach based on numerical simulations that allows measurement
363 error to be included in the model. Contrary to common practice, this approach ensures that
364 practitioners can control the proportion of false negatives when identifying a contaminated
365 value in a sample. This method is then extended to sample pooling, where the threshold is
366 corrected for measurement error and pool size. From a practical point of view, the approach

Journal Pre-proofs

367 also allows to derive a threshold value above which a sample or a pool should be declared as
368 contaminated directly on the basis of the measurement, without any additional adjustment or
369 correction, since the measurement error has already been taken into account by the model in
370 the derivation of the threshold value. Even if in one of the case studies the sample pooling
371 turned out not to be relevant, in most case studies, it enabled sensible gain in terms of the
372 relative number of samples to be tested.

373 Acknowledgments

374 This work was supported by the European Union's Horizon 2020 Research and Innovation
375 Program under Grant Agreement No 861917 – SAFFI as well as by the National Program on
376 Key Research Project of China (2019YFE0103900). The authors would like to thank Dr
377 Vasiliki Giatrakou of Yiotis Anonimos Emporiki & Viomixaniki Etaireia (Athens, Greece) for
378 his assistance with enzyme-linked immunosorbent assays (ELISA).

Journal Pre-proofs

379 # References

380 Aldridge, M., Johnson, O., & Scarlett, J. (2019). Group Testing : An InformationTheory
381 Perspective. *Foundations and Trends® in Communications and Information Theory*,
382 *15*(3-4), 196-392. https://doi.org/10.1561/0100000099

383 Battilani, P., Palumbo, R., Giorni, P., Dall'Asta, C., Dellafiora, L., Gkrillas, A., Toscano, P.,
384 Crisci, A., Brera, C., De Santis, B., Rosanna Cammarano, R., Della Seta, M., Campbell,
385 K., Elliot, C., Venancio, A., Lima, N., Gonçalves, A., Terciolo, C., & Oswald, I. P. (2020).
386 Mycotoxin mixtures in food and feed : Holistic, innovative, flexible risk assessment
387 modelling approach: *EFSA Supporting Publications*, *17*(1).
388 https://doi.org/10.2903/sp.efsa.2020.EN-1757

389 Berninger, J. P., & Tillitt, D. E. (2019). Polychlorinated biphenyl tissue-concentration thresholds
390 for survival, growth, and reproduction in fish. *Environmental Toxicology and Chemistry*,
391 *38*(4), 712-736. https://doi.org/10.1002/etc.4335

392 Bilder, C. R., Iwen, P. C., Abdalhamid, B., Tebbs, J. M., & McMahan, C. S. (2020). Tests in
393 Short Supply? Try Group Testing. *Significance*, *17*(3), 15-16.
394 https://doi.org/10.1111/1740-9713.01399

395 Brodal, G., Aamot, H. U., Almvik, M., & Hofgaard, I. S. (2020). Removal of Small Kernels
396 Reduces the Content of Fusarium Mycotoxins in Oat Grain. *Toxins*, *12*(5), 346.
397 https://doi.org/10.3390/toxins12050346

398 Cai, P., Liu, S., Tu, Y., & Shan, T. (2023). Toxicity, biodegradation, and nutritional intervention
399 mechanism of zearalenone. *Science of The Total Environment*, 168648.
400 https://doi.org/10.1016/j.scitotenv.2023.168648

401 Cheng, X., Chavez, R. A., & Stasiewicz, M. J. (2020). When to use one-dimensional, two-
402 dimensional, and Shifted Transversal Design pooling in mycotoxin screening. *PLOS ONE*,
403 *15*(8), e0236668. https://doi.org/10.1371/journal.pone.0236668

404 Currie, M. J., McNiven, M., Yee, T., Schiemer, U., & Bowden, F. J. (2004). Pooling of Clinical
405 Specimens Prior to Testing for *Chlamydia trachomatis* by PCR Is Accurate and Cost
406 Saving. *Journal of Clinical Microbiology*, *42*(10), 4866-4867.
407 https://doi.org/10.1128/JCM.42.10.4866-4867.2004

408 Dorfman, R. (1943). The Detection of Defective Members of Large Populations. *The Annals of*
409 *Mathematical Statistics*, *14*(4), 436-440. https://doi.org/10.1214/aoms/1177731363

410 Engel, E., Rivière, G., Kemmer, D., Deusch, O., Fuchsbauer, N., Biesterveld, S., Krystalli, E.,
411 Bondoux, M., Li, G., Yang, W., Hou, J., Liang, Y., Yang, H., Fang, W., Pettoello-
412 Mantovani, M., Flynn, B., Rantsiou, K., Van Der Burg, B., Bover-Cid, S., & Zwietering,
413 M. H. (2022). Safe food for infants : An EU-China project to enhance the control of safety
414 risks raised by microbial and chemical hazards all along the infant food chains. *Global*
415 *Pediatrics*, *2*, 100009. https://doi.org/10.1016/j.gpeds.2022.100009

416 Eskola, M., Kos, G., Elliott, C. T., Hajšlová, J., Mayar, S., & Krska, R. (2020). Worldwide
417 contamination of food-crops with mycotoxins : Validity of the widely cited 'FAO

418 estimate' of 25%. *Critical Reviews in Food Science and Nutrition*, *60*(16), 2773-2789.
419 https://doi.org/10.1080/10408398.2019.1658570

420 Giray, B., Atasayar, S., & Sahin, G. (2009). Determinatıon of ochratoxin A and total aflatoxin
421 levels in corn samples from Turkey by enzyme-linked immunosorbent assay. *Mycotoxin*
422 *Research*, *25*(2), 113-116. https://doi.org/10.1007/s12550-009-0016-0

423 Hann, M., Hudson, B., Lewell, X., Lifely, R., Miller, L., & Ramsden, N. (1999). Strategic
424 Pooling of Compounds for High-Throughput Screening. *Journal of Chemical Information*
425 *and Computer Sciences*, *39*(5), 897-902. https://doi.org/10.1021/ci990423o

426 Jagadeesan, B., Bastic Schmid, V., Kupski, B., McMahon, W., & Klijn, A. (2019). Detection of
427 Listeria spp. And L. monocytogenes in pooled test portion samples of processed dairy
428 products. *International Journal of Food Microbiology*, *289*, 30-39.
429 https://doi.org/10.1016/j.ijfoodmicro.2018.08.017

430 Köppen, R., Koch, M., Siegel, D., Merkel, S., Maul, R., & Nehls, I. (2010). Determination of
431 mycotoxins in foods : Current state of analytical methods and limitations. *Applied*
432 *Microbiology and Biotechnology*, *86*(6), 1595-1612. https://doi.org/10.1007/s00253-010-
433 2535-1

434 Mallapaty, S. (2020). The mathematical strategy that could transform coronavirus testing.
435 *Nature*, *583*(7817), 504-505. https://doi.org/10.1038/d41586-020-02053-6

436 Merritt, F. D. (1898). Cours d'Economie Politique. Vilfredo Pareto. *Journal of Political*
437 *Economy*, 6(4), 549-552. https://doi.org/10.1086/250536

438 Pereira, C., Cunha, S. C., & Fernandes, J. O. (2022). Mycotoxins of concern in children and
439 infant cereal food at European level: Incidence and bioaccessibility. *Toxins*, 14(7), 488.
440 https://doi.org/10.3390/toxins14070488

441 Ratel, J., Vigneau, E., Mercier, F., Lebreton, S., Nguyen, T. L. N., Courcoux, P., & Engel, E.
442 (2024). Implementation of sample pooling to strengthen the surveillance of food chemical
443 safety : Case study of nDL-PCBs in pork meat. *Food Chemistry*, *433*, 137111.
444 https://doi.org/10.1016/j.foodchem.2023.137111

445 Trevisi, P., Luise, D., Spinelli, E., Correa, F., De Leo, E., Trambajolo, G., Diegoli, G., & Bosi,
446 P. (2020). Transfer of Mycotoxins from Lactation Feed to Colostrum of Sows. *Animals*,
447 *10*(12), 2253. https://doi.org/10.3390/ani10122253

448 Verdun, C. M., Fuchs, T., Harar, P., Elbrächter, D., Fischer, D. S., Berner, J., Grohs, P., Theis,
449 F. J., & Krahmer, F. (2021). Group Testing for SARS-CoV-2 Allows for Up to 10-Fold
450 Efficiency Increase Across Realistic Scenarios and Testing Strategies. *Frontiers in Public*
451 *Health*, *9*, 583377. https://doi.org/10.3389/fpubh.2021.583377

452 Wan, J., Chen, B., & Rao, J. (2020). Occurrence and preventive strategies to control mycotoxins
453 in cereal-based food. *Comprehensive Reviews in Food Science and Food Safety*, 19(3),
454 928-953. https://doi.org/10.1111/1541-4337.12546

455 Wei, W., Liu, C., Ke, P., Chen, X., Zhou, T., Xu, J., & Zhou, Y. (2021). Toxicological and
456 physiological effects of successive exposure to ochratoxin A at food regulatory

Journal Pre-proofs

457 limits. *Food and Chemical Toxicology*, 151, 112128.
458 https://doi.org/10.1016/j.fct.2021.112128

459 Więckowska, M., Cichon, N., Szelenberger, R., Gorniak, L., & Bijak, M. (2024). Ochratoxin A
460 and Its Role in Cancer Development: A Comprehensive Review. *Cancers*, 16(20), 3473.
461 https://doi.org/10.3390/cancers16203473

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480 # Tables

481 **Table1**.

482 Method performance for ZEA and OTA determination.

case	LOQ [1] (μg/kg)	pLOQ[2]	ML[3]	Prevalence[4]	RSD[5]

Journal Pre-proofs

			(μg/kg)			
	I	10	66%	20	8%	17.4%
ZEA	II	10	61%	50	0.2%	8%
	III	10	74%	100	<0.1%	5%
OTA		0.3	50%	3	<0.1%	10%

483 [1]LOQ: Limit of Quantification

484 [2]pLOQ: the proportion of values below LOQ

485 [3]ML: Maximum Limit (as described in Regulation (EU) 2023/915)

486 [4]Prevalence: proportion of observations exceeding ML.

487 [5]RSD: Relative Standard Deviation

488

489

490

491

492

493

494

495

496

497

498

499

500

501

Journal Pre-proofs

502 **Table 2**:

503 Corrected Maximum Limit, ML_p^* as a function of the pool size for the different case studies.

	ZEA I	ZEA II	ZEA III	OTA
	ML=20, LOQ=10, RSD=17.5%,prev=8%	ML=50, LOQ=10, RSD=8%,prev=0.2%	ML=100, LOQ=10, RSD=5%,prev=0.1%	ML=3, LOQ=0.3, RSD=10%,prev=0.1%
n_P	ML_P^*	ML_P^*	ML_P^*	ML_P^*
1	15	44.7	95	2.6
2	9.1	25.5	50.1	1.4
3	7.7	19.5	35.7	1
4	7.1	16.7	28.7	0.8
5	6.7	15	24.4	0.7
6	6.5	13.8	21.5	0.6
7	6.2	13	19.5	0.6
8	6.1	12.4	18	0.5
9	6	11.9	16.9	0.5
10	5.9	11.6	16	0.5
20	5.7	9.9	11.9	0.4
30	5.6	9.4	10.5	0.3
40	5.7	9.2	9.9	0.3

Journal Pre-proofs

50	5.8	9.1	9.6	0.3

504 *When ML_p^* are larger than the LOQ, values are in bold.*

505

506

507

508

509

510

511

512

513

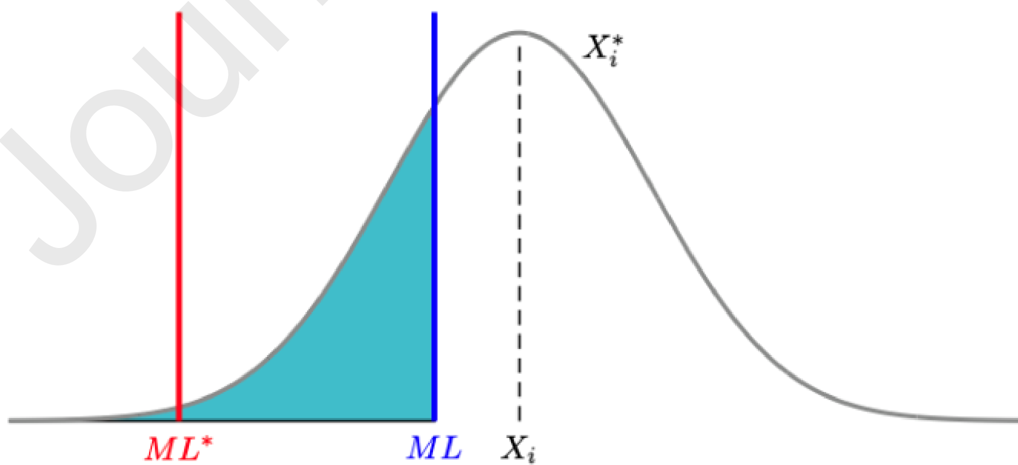
514

515

516

517 # Figures

518 **Figure 1.**



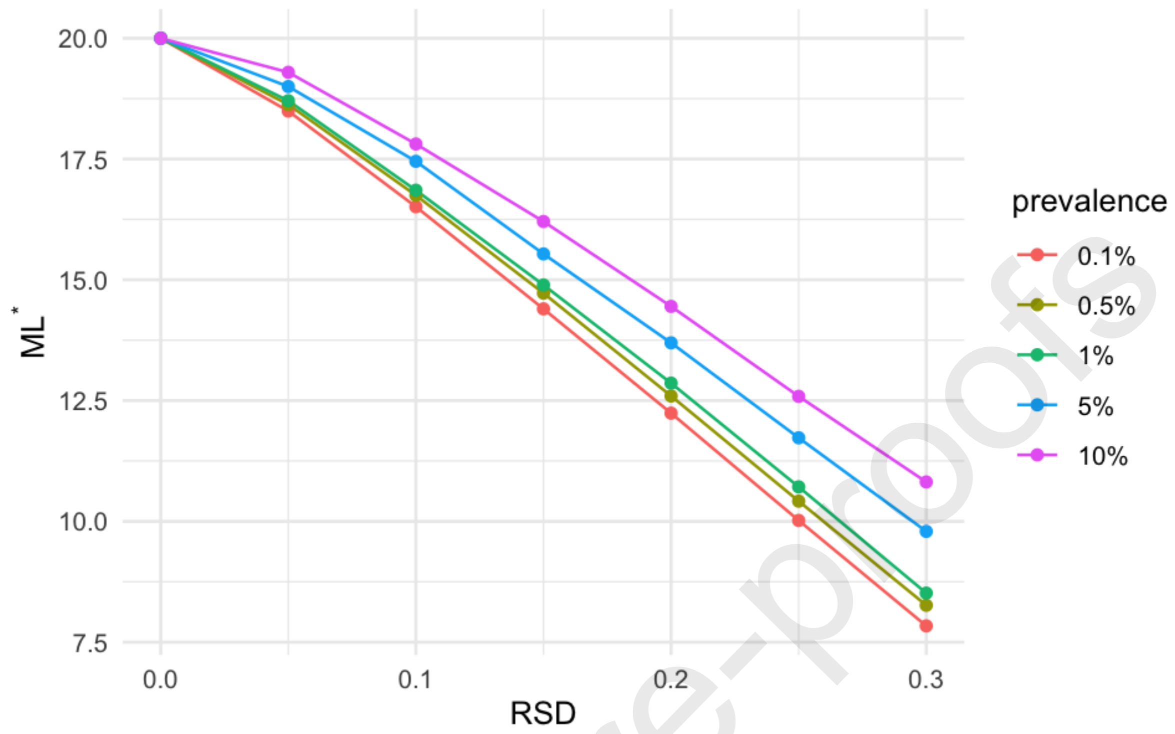
519

520

Journal Pre-proofs

521 **Figure 2.**

522

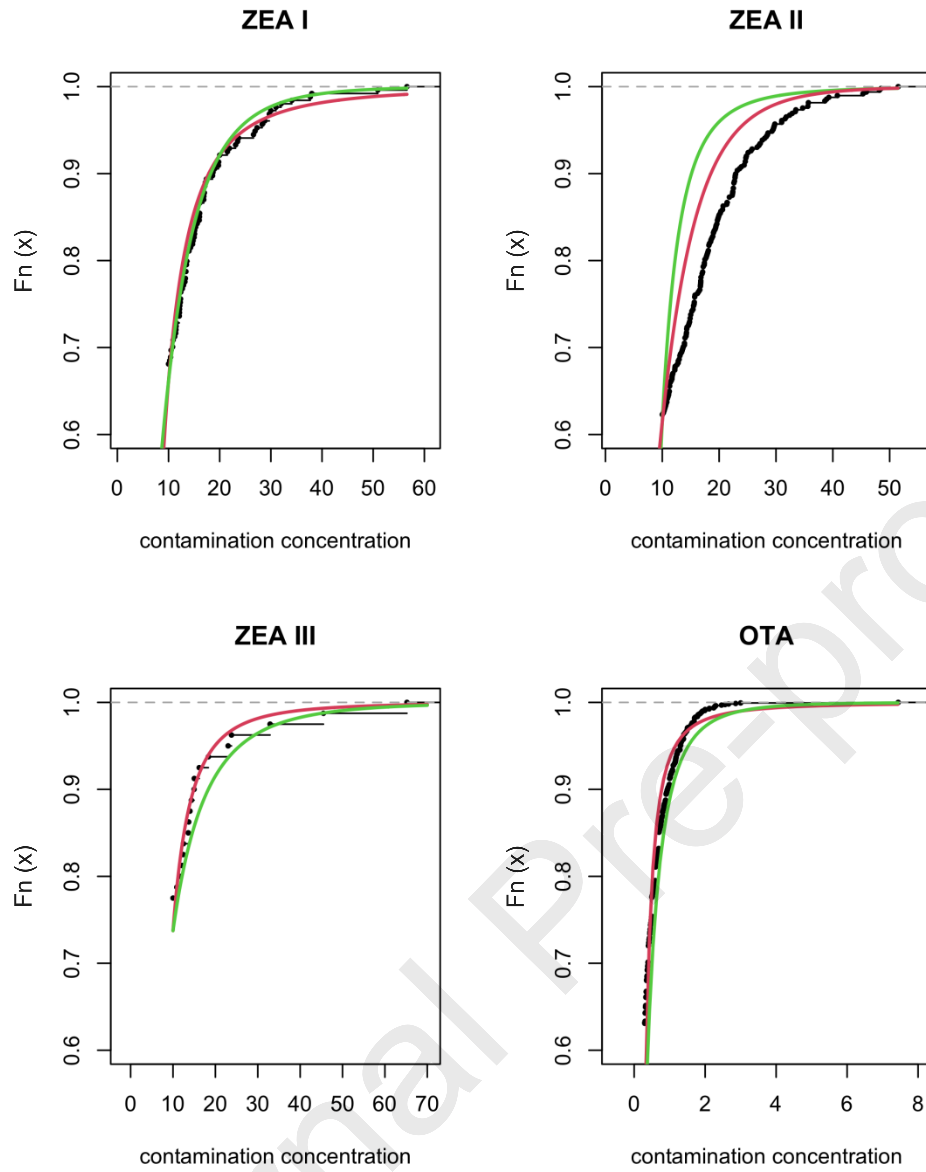


523

524

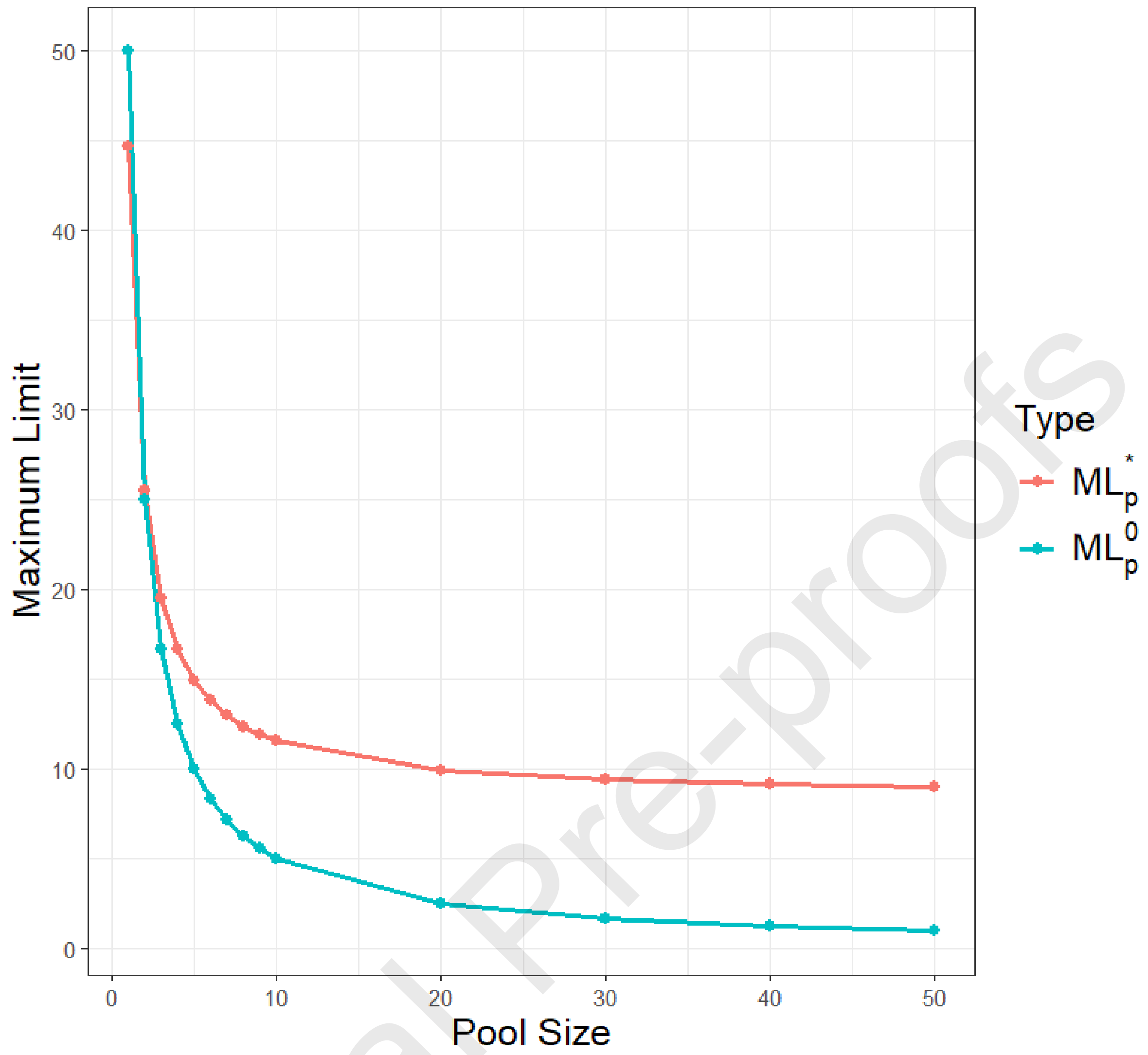
525 **Figure 3.**

Journal Pre-proofs



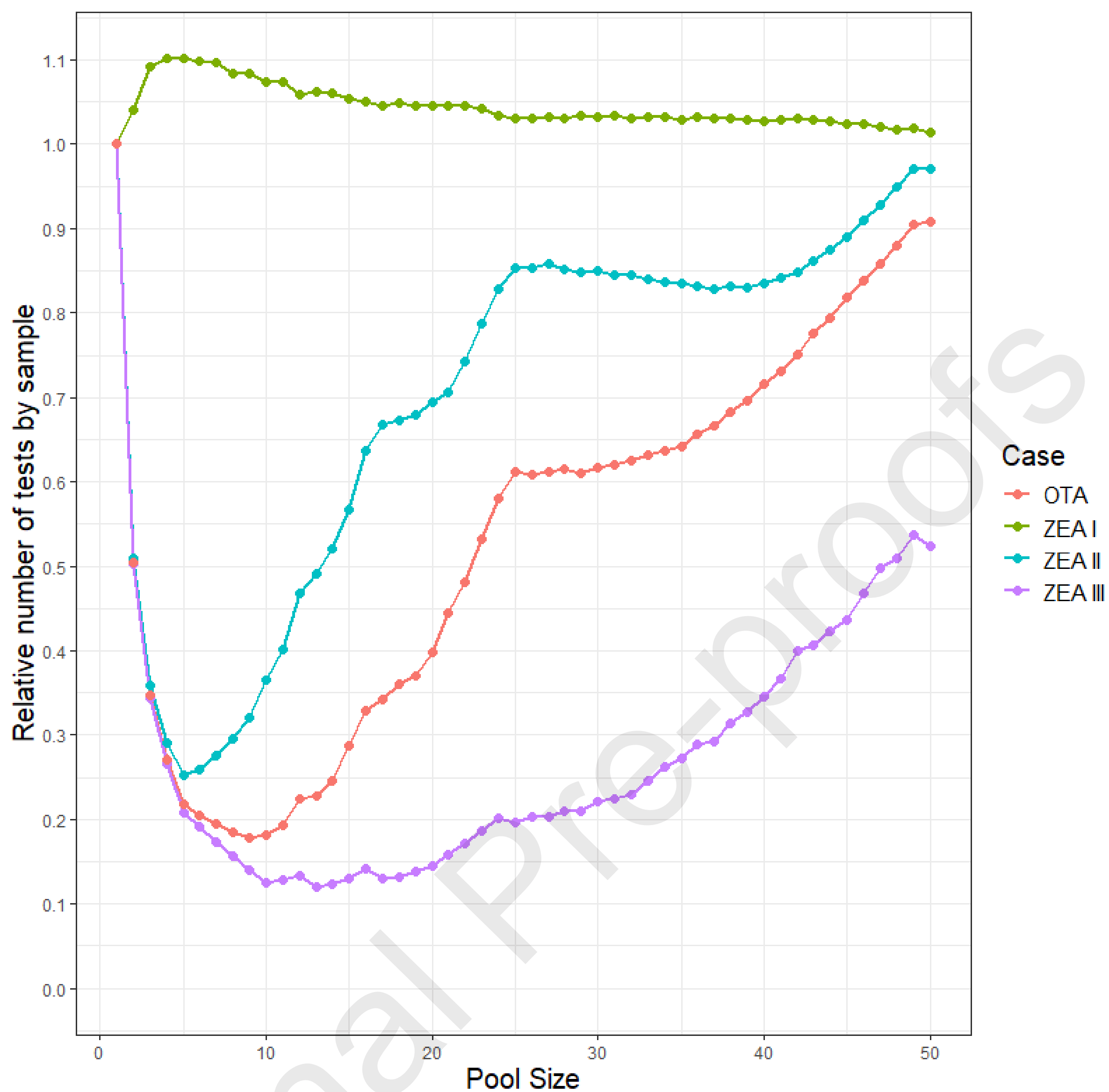
526
527
528
529
530
531
532
533
534 **Figure 4.**

Journal Pre-proofs



535
536
537
538
539
540
541
542
543
544
545 **Figure 5.**

Journal Pre-proofs



546

547

548

549

550

551

552

553

554

555 **Figure captions**

Journal Pre-proofs

556 **Figure 1**. Distribution of a measurement X_i^* according to a true value X_i closed to ML. The area
557 in blue corresponds to proportion of measurements below the ML. ML^* designates the corrected
558 ML (i.e. the threshold calculated considering measurement error).

559 **Figure 2.** Evolution of the corrected Maximum Limit, ML^*, as a function of RSD and the
560 prevalence, for a sensitivity target of 99% and ML=20.

561 **Figure 3.** Empirical Cumulative Distribution function of the data and Cumulative Distribution
562 Function of the corresponding Pareto distribution (red line) and Log-normal distribution (green
563 line) for each of the four case studies (ZEA I, ZEA II, ZEA III, OTA).

564 **Figure 4.** ML_P^* and ML_P^0 as a function of the pool size considering the parameters of ZEA II.

565 **Figure 5.** Relative number of tests by sample as a function of the pool size for the different
566 case studies considering 50 samples in the analytical batches, and the corrected Maximum
567 Limits for decision making.

568
Highlights

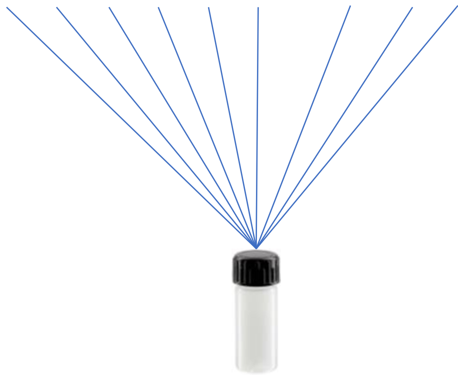
569 - Sample pooling was applied to mycotoxin self-monitoring in cereal-based foods
570 - Mycotoxin concentrations were fitted by Pareto distributions
571 - A numerical simulation of pooling was implemented using the Dorfman-2-step strategy
572 - Pooling strategy can decrease the number of analysis per sample by more than four
573 - The tests' sensitivity and specificity were not impaired with pooling in most cases

574

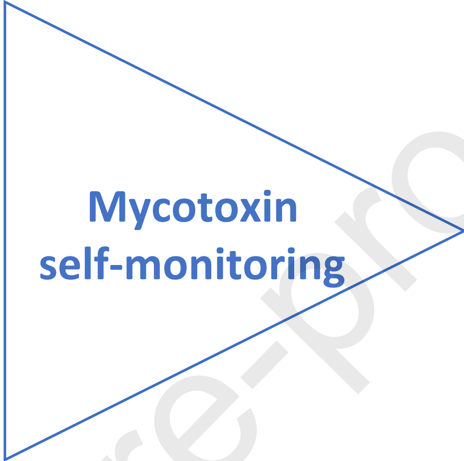
575

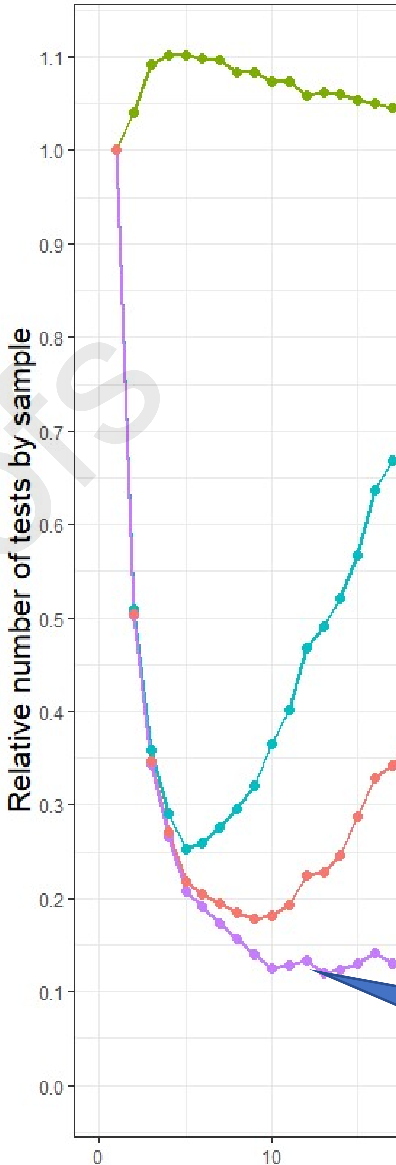
Journal Pre-proofs

Cereal samples



Sample pooling





576

577

Journal Pre-proofs