Food Control 22 (2011) 1085–1094

Contents lists available at ScienceDirect

Food Control

journal homepage: www.elsevier.com/locate/foodcont

Application of binomial and multinomial probability statistics to the sampling design process of a global grain tracing and recall system☆

Kyung-Min Lee [a], Paul R. Armstrong [b], J. Alex Thomasson [c], Ruixiu Sui [d], Mark Casada [b], Timothy J. Herrman [a,*]

[a] Office of the Texas State Chemist, Texas Agricultural Experiment Station, Texas A&M University System, College Station, TX 77841, United States
[b] Center for Grain and Animal Health Research, Agricultural Research Service, U.S. Department of Agriculture, Manhattan, KS 66502, United States
[c] Biological and Agricultural Engineering Department, Texas A&M University, College Station, TX 77843, United States
[d] Cotton Ginning Research Unit, Agricultural Research Service, US Department of Agriculture, Stoneville, MS 38776, United States

A R T I C L E I N F O

Article history:
Received 14 July 2010
Received in revised form
15 December 2010
Accepted 22 December 2010

Keywords:
Grain traceability system
Sampling design process
Binomial probability distribution
Multinomial probability distribution
Sample size

A B S T R A C T

Small, coded, pill sized tracers embedded in grains are proposed as a method to store a historical record of grains and retrieve coded information for grain traceability. This study aimed to develop and validate a statistical sampling procedure to securely collect sample sizes (kg) and number of tracers since the sampling accuracy is critical in the proposed traceability system for capturing information and data related to grain lots to trace the grain back through the route in a grain supply chain. The statistical results and observations showed similar concentrations and insignificant segregation of tracers in bin and truck operations. The number of tracers required for identification of grain sources fell within the confidence intervals and sample sizes (kg) estimated by statistical probability methods. Truck sampling appeared more feasible in collecting the secure number of tracers over bin sampling. The designed sampling process was empirically proven to be practically applicable and provide better scientific assurance of sampling accuracy, which may reduce economic risks and their consequent costs caused by unfavorable sampling in the propose traceability system.

1. Introduction

The importance and efforts of developing a traceability system for the grain industry have significantly increased during the past decade in several countries due to concerns and issues relating to biotechnology products, food safety regulations, bio-security policy, and the adoption of the quality management system (Thakur & Hurburgh, 2009). The European Commission (EC) of the European Union (EU) uses traceability systems to track agricultural products produced in Europe and imported from foreign countries (Official Journal of the European Communities, 2002), having enacted laws which provide authorities the power to investigate and recall unsafe products distributed within the market. The United States (US) grain traceability system tends to be developed mainly due to economic incentives (Golan et al., 2004; USDA, 2005); the rule for Establishment and Maintenance of Records for Implementing Section 306 of the Bio-terrorism Act of 2002 (US Congress, 2002) requires grain tracing to be one step forward and backward (Federal Register, 2004). In Canada, Can-Trace was established in 2003 to develop traceability standards for all food products in Canada, and it is based on a one-up/one-down model of sharing traceability information (Can-trace, 2003). The International Standards Organization (ISO) offered a new food safety standard (ISO 22005, 2007) on traceability with general principles and requirements including the responsibility of all system participants to provide, on demand, all necessary records and product information to authorities.

A typical grain supply chain in the US involves several steps handling a crop, starting from a seed company that supplies seeds to farmers. Grain harvested and stored on the farm move through country elevators, inland sub-terminals, processing facilities, and export terminals by two or more transportation modes, depending on origin and final destination in the grain supply chain (Herrman, 2002; Meyer, 2004). Grain transportation modes between origin and final destination are determined by accessibility, costs, local requirements, and marketability (Meyer, 2004). Grain movement to

☆ Mention of trademark or proprietary products does not constitute a guarantee or warranty of the product by the U.S. Department of Agriculture and does not imply its approval to the exclusion of other products that may also be suitable.

* Corresponding author. Tel.: +1 979 845 1121; fax: +1 979 845 1389.
E-mail address: tjh@otsc.tamu.edu (T.J. Herrman).

0956-7135/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.foodcont.2010.12.016

elevators or processing facilities means the entry of grains into the commercial grain market. Grain elevators serve as collection points and an important quality control point for grain. Usually, trucks are weighed, sampled for grain quality evaluation, and unloaded at the dump pit. The delivered grains at elevators are cleaned, sorted, and blended to make a homogeneous grain quality and to meet defined quality standards. Large processing facilities, such as wet and dry mills receive whole grains from many different geographical regions and via a large number of truck and railcar shipments (Herrman, 2002; Sims, 2005). Grains produced in geographical regions far away from export terminals access to the export terminals via barges, railcars, and trucks. At an export terminal, a commodity export vessel may contain commingled grains from the tens of thousands of individual truckloads or sources (Sims, 2005).

Traceability systems for grains should be feasible to trace their origin throughout the supply chain and need to be developed by considering all factors associated with the market place to address emerging issues and to keep competitiveness for producers and handlers at the marketing system (Golan et al., 2004). Identifying grain origin becomes complicated by blending grains with different origins at local elevators and several additional changes and combinations occurred at the market place. The type of product and objectives of the traceability systems that firms want to achieve will dictate the depth and precision of the systems at the market place. The traceability systems would work more efficiently as they are coupled with other management, marketing, safety and quality control tools to accomplish the goals of the system's users. However, instead of the standard system, the traceability systems developed on the diverse markets and products may result in an increase in service times and costs for transportation and marketing costs throughout the supply chain depending on crops and traceability requirements among countries (Herrman, 2002).

Obtaining a truly representative grain sample is important to obtain accurate information on the grain population for determining grain quality characteristics and values. An adequate sampling knowledge and understanding are also critical in determining the sampling method, sampling strategy, and sample size (Hellevang, Backer, & Maher, 1992). A grain sampling protocol should be well-defined and accurate since a biased sample invalidates analytical results and hinders determining the appropriate action for grain quality management (Coker et al., 1995). An appropriate sampling protocol is particularly important when grains and other components (e.g., broken kernels and foreign materials) in transport or storage are stratified and segregated, which increases the variation in the physicochemical properties within a container. Increasing a sample size improves the accuracy of estimates and narrows the width of a confidence interval. The determination of the number of samples, sampling location, and sampling pattern is influenced by the sampling device and the distribution of target objectives and analytes in the grain (Hellevang et al., 1992).

The current private traceability and identity preservation (IP) systems for grains can provide the customer with a higher-value product, improve quality grain, minimize damage and theft problems, prevent seasonal fluctuations, better manage and handle products with limited marketability, and build brand loyalty, which may give substantial benefits for producers and companies (ERS, 2001; Herrman, 2002; USDA, 2005; Vachal & Tolliver, 2001). However, current existing traceability and IP systems are expensive and labor-intensive process compared to the bulk grain delivery system (Vachal & Tolliver, 2001). If the systems fail to segregate biotech products and specialized grains, the products would be commingled and consumers' confidence on the systems would be lost (ERS, 2001). On the contrary, the proposed tracer-based system proposed in this study can be characterized with low cost and easy implementation for the current grain transportation and processing systems in the U.S. as well as in other countries. Mutual benefits may also be derived when combined with existing traceability and recall systems. Coding the field or farm location on the tracers would enable tracing system participants and customers to trace the grain movement even after grains are commingled and shipped to multiple distributors. The U.S. grain handling infrastructure moves large volumes of grain from multiple sources to end-users or export grain elevators. In view of this complex system, a need exists to design a scientific sampling process that verifies the grain origin as it moves through this complex handling system. Such a sampling process requires the application of probability statistics for the tracer-based system.

The objectives of this study were to design the statistical sampling procedure to collect the secure number of tracers for identification of grain sources in a grain lot, investigate the variance structure of tracer concentration at different sampling points, and understand the difference in tracer concentration and distribution between theoretically expected samples and actually collected samples. The science-based sampling design process would help to improve the practical sampling accuracy and facilitate the implementation of the new conceptual traceability system.

2. Materials and methods

2.1. Delineation of the proposed traceability system

Fig. 1 shows the general flow of grain and possible sampling points throughout the proposed traceability system. The first or second sample collection and grain identification point in this conceptualized tracing system involves truckloads of grains delivered to country elevators, where representative probe samples of grain would be collected prior to the receiving pit. At inland and river or export terminals which receive the grain from country elevators with multiple grain sources, capturing tracers may require sample collection from a grain transport container, through stream sampling at the receiving pit, or through automated sampling at the bucket elevator. Note that the complete removal of the tracers, prior to end-use, is not necessary due to the use of food grade ingredients to manufacture tracers.

2.2. Probability theory

In designing the sampling process, the application of probability statistics was considered to provide a science-based sampling approach for the better estimation of the tracer distribution in bulk grain. Tracers mixed with grain are conceptualized in Fig. 2 along with the probability variables. In binomial probability statistics, which are applicable for sampling grain lots containing tracers and single-source grains, the mean tracer concentration (μ) and a standard deviation (σ) are computed as $\mu = np$ and $\sigma = \sqrt{np(1-p)}$, respectively, where n is the number of independent trials or samples and p denotes the probability of success in an individual trial (observed sample portion). A sampling system that identifies commingled tracers and grains from different origins of fields or sources at elevators and terminals in the grain supply chain requires the application of multinomial probability statistics. In a multinomial distribution, which is an extension of the binomial distribution, each event falls into one of k possible outcomes with respective probabilities $p_1,\ldots,p_k$ (Larsen & Marx, 1985). The probability of a multinomial distribution with the random variables X_i is defined by:

$$P(X_1 = x_1, \cdots, X_k = x_k) = \frac{n!}{x_1!\ldots\ldots x_k!} p_1^{x_1} \cdots p_k^{x_k}$$

where x_i are non-negative integers with $\sum_{i=1}^{k} x_i = n$, and p_i are constants with $\sum_{i=1}^{k} p_i = 1$. The mean (μ_i) and variance (σ_i^2) of the X_i are $\mu_i = np_i$ and $\sigma_i^2 = np_i(1-p_i)$, respectively.

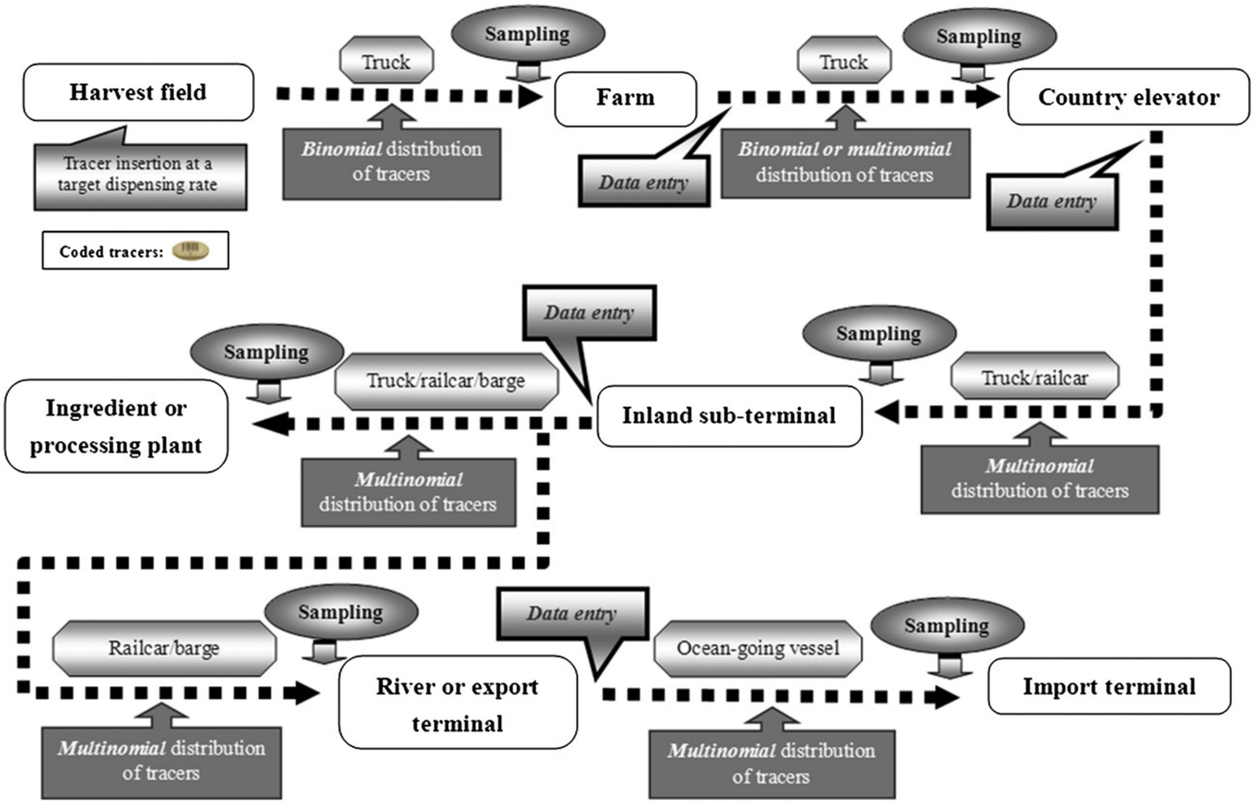

Fig. 1. General process flow diagram for the proposed grain tracing and recall system.

For the application of the probability statistics to practical sampling, the grains and tracers delivered to an individual container during grain transporting and storage were considered as random variables X_i (Fig. 2). Likewise, the total sample size and probability of taking tracers falling in the ith container correspond to n independent trials and the probability of X_i outcomes, respectively. The p_i denotes the proportion of tracers for the ith container. A computed sample size (n) was considered as the total number of grain kernels and tracers in a sample. The number of kernels and tracers were converted into a sample mass (kg) based on the wheat thousand kernel weight (28.1 g) for ease of computation and users' convenience.

Three binomial and four multinomial probability methods were evaluated for application to this traceability system and were used to construct confidence intervals for the tracer distribution (Table 1). For the construction of a binomial confidence interval, the Wald interval (1943), the Wilson interval (1927), and the Agresti-Coull interval (1998) were used and compared. The binomial confidence intervals provide a basic formula to determine a sample size (kg) for the identification of grains as they are hauled from a single farm to the country grain elevator. Likewise, simultaneous confidence intervals for multinomial proportions of tracers were estimated and compared using four methods, including Fitzpatrick and Scott method (1987), Goodman method (1965), Quesenberry and Hurst method (1964), and a one-at-a-time binomial interval method.

The methods for constructing simultaneous confidence intervals have been employed by several authors (Angers, 1984; Thompson,

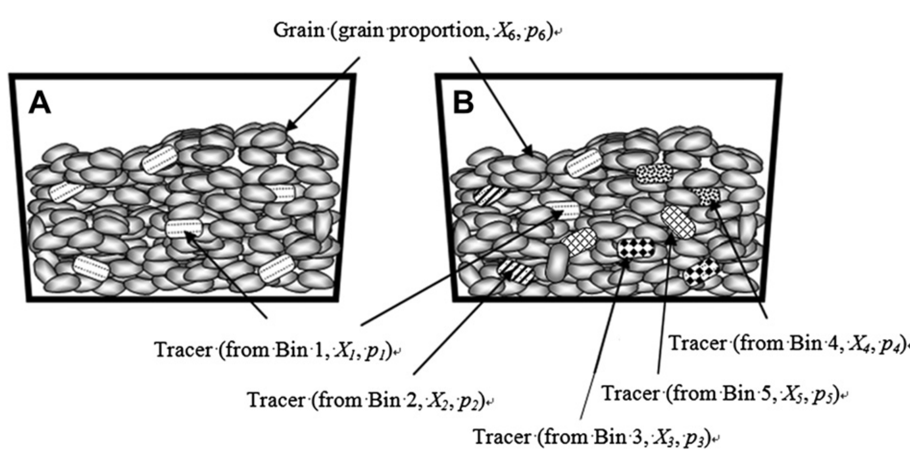

Fig. 2. Binomial (A) and multinomial distribution (B) of commingled tracers and grains in a container. The X_i denotes a probability variable for grains or tracers dispensed at the ith container. The p_i is the proportion of grains or tracers for the ith container.

Table 1
Statistical probability methods for estimating confidence intervals and methods for calculating sample size[a].

Methods	Formulas
Binomial probability methods	
1) Wald interval (1943)	$\widehat{p} \pm z_{\alpha/2}\sqrt{\widehat{p}(1-\widehat{p})/n}$
2) Wilson interval (1927)	$\dfrac{\widehat{p}+\frac{z^2_{\alpha/2}}{2n} \pm z_{\alpha/2}\sqrt{\frac{\widehat{p}(1-\widehat{p})}{n}+\frac{z^2_{\alpha/2}}{4n^2}}}{1+\frac{z^2_{\alpha/2}}{n}}$
3) Agresti-Coull interval (1998)	$\tilde{p} \pm z_{\alpha/2}\sqrt{\tilde{p}(1-\tilde{p})/(n+z^2_{\alpha/2})}$
Multinomial probability methods	
1) Quesenberry and Hurst method (1964)	$p_i = \widehat{p}_i \pm \chi_{\alpha,k-1}[\widehat{p}_i(1-\widehat{p}_i)/n]^{1/2}$, $i = 1,2,\ldots\ldots k$
2) Goodman method (1965)	$p_i = \widehat{p}_i \pm \chi_{\alpha/k,1}[\widehat{p}_i(1-\widehat{p}_i)/n]^{1/2}$, $i = 1,2,\ldots\ldots k$
3) Fitzpatrick and Scott method (1987)	$p_i = \widehat{p}_i \pm \frac{z(\alpha/2)}{2\sqrt{n}}$, $i = 1,2,\ldots\ldots k$
Methods for sample size	
1) Angers method (1984)	$n = \min\{\max[z^2(\alpha_{i/2})/4d_i^2; i = 1,2,\ldots\ldots,k]\}$
2) Tortora method (1978)	$n = Kp_i(1-p_i)/d_i^2$
3) Thompson method (1987)	$n = \max_m z^2(1/V)(1-1/V)/d^2$

$\widehat{p}$ = observed sample proportion.
$z_{\alpha/2} = (1-\alpha/2)$100th percentile of the standard normal distribution.
n = the total sample size.
$\tilde{p} = \dfrac{T+z^2_{\alpha/2}/2}{n+z^2_{\alpha/2}}$ (T = observed sample frequency).
$\chi_{\alpha,k-1}$ = upper $(1-\alpha)$100th percentile of the chi-square distribution with k-1 degree of freedom
$\chi_{\alpha/k,1}$ = upper $(1-\alpha/k)$100th percentile of the chi-square distribution with 1 degree of freedom
$z(\alpha)$ is the upper $(1-\alpha)$100th percentile of the standard normal distribution
α_i = significance level of the ith confidence interval, $\sum_{i=1}^{k}\widehat{\alpha}_i = \widehat{\alpha}$ $(0 < \alpha < 1)$
d_i = half-width of the ith confidence interval
$K = (1-\alpha/k)$100th percentile of the chi-square distribution with 1 degree of freedom
V = number of outcomes with the non-zero proportion ($P \neq 0$) in the population.
[a] Symbols in formulas.

1987; Tortora, 1978) for choosing a sample size for multinomial distribution (Table 1). The computational procedures for these methods were executed using Microsoft Excel (Microsoft Corp., Redmond, WA). The half lengths (d_i, precision value) of the confidence intervals were adjustable in those methods so that the sample size could be calculated based on desired levels of precision for particular tracer concentration and distribution. This feature can be applied to optimize the sample size in practice when the predetermined sampling plan is believed to be unfit for a particular sampling point. If parameter $\widehat{p}_i$ is unknown in practice, then $\widehat{p}_i$, a least favorable value (worst case), of the overall mean tracer concentration at each sampling point may be used, which inevitably would result in redundant samples.

2.3. Tracer preparation

White-colored tracers were made of 97% processed sugar (Di-Pac Direct Compacting & Tableting Sugar, Domino Foods, Inc., Baltimore, MD) and 3% magnesium stearate (TABLETpress.NET, Athens, OH), while the other four colored tracers were manufactured by substituting 2% of the processed sugar with four different dye powders. The pre-mixed formulations were directly compressed with a tablet press (FSP-30 Single Punch Tablet Press, Minhua, Shanghai, China) using the tracer-forming die with dimensions of 11 mm in length × 6 mm in width × 4 mm in height at a production rate of 50 tracers/min.

2.4. Probability sampling plan and sampling units

Probability sampling tests were conducted with the five colored tracers to validate the sampling design process in the research elevator and a large outdoor bin at the USDA-ARS, Center for Grain and Animal Health Research in Manhattan, Kansas. Prior to the tests, the truck, conveyor belts, bins, and spouts were thoroughly cleaned to remove any residuals and contaminants from previous loading operations. Thereafter, the wheat stored in the elevator bins for the tests was moved down to the receiving pit and bucket-elevated to the top machinery floor. Targeting a tracer concentration of three tracers

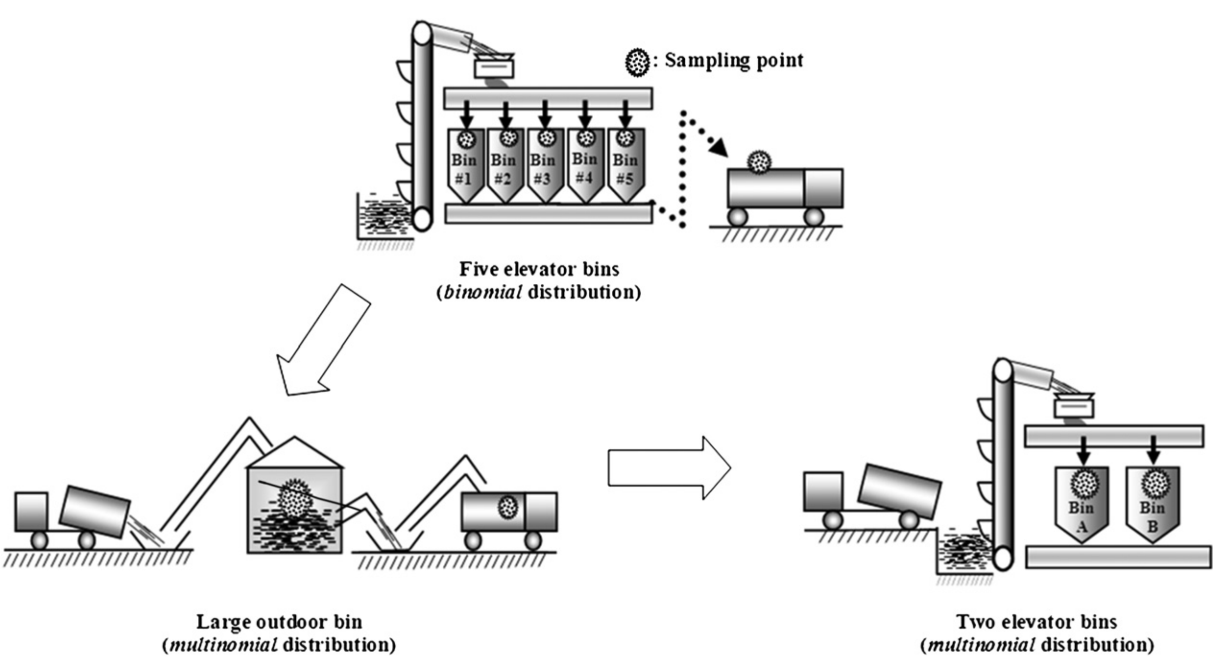

Fig. 3. Grain flow and sampling points for the probability sampling tests in the downscaled grain traceability system.

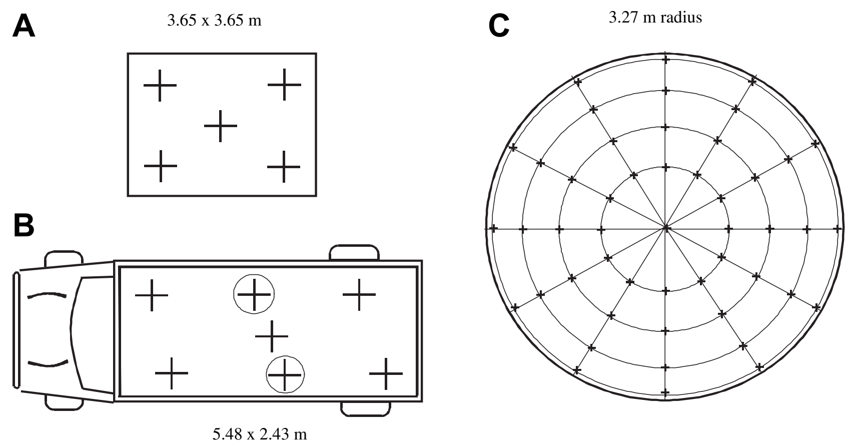

Fig. 4. Sampling patterns in an elevator bin (A), a truckload (B), and a large outdoor bin (C). The + in each container represents a sampling location. The ⊕ location in a truckload was not sampled during the binomial probability test.

per kg of wheat, tracers were dispensed into constant grain flowing at the rate of 13.7 t h^{-1} (506 bu hr^{-1}) using a tracer dispenser. Wheat blended with single-colored tracers discharged to a weigh-scale on the top floor of the elevator. The weigh-scale automatically dumped accumulated grain in 408-kg increments. Approximately 11.6 kg of tracers were used for each tracer color, resulting in a final tracer concentration averaging 2.93 tracers per kg of wheat for all of the five colored tracers. Bins containing single-colored tracers were assumed as a single grain source.

Tracer concentration and distribution were examined at the predefined sampling points, shown in Fig. 3, simulating supply chain traceability during grain transportation and storage. Grain stored in each elevator bin was loaded onto a truck with a capacity of 8.1 tons (300 bu) and moved to a large (136,000 kg) outdoor bin. This grain movement simulated the aggregation of grain coming from several storage bins; individual lots lost identity by combining grain from the different lots. Grain stored in the outdoor bin was then loaded onto the same truck after the bin sampling, and transferred back to two elevator bins (1200 bu each). The truck was sampled prior to dumping in the receiving pit. In a commercial grain elevator, an incoming grain lot purchased from a single seller is often separated into several storage bins. All of the grain samples were taken with a 4-ft probe from the upper bin level of the sampling locations as deep as possible while truck probing was performed evenly over the entire bed and at the bottom of the bed using the systematic sampling method (Fig. 4). For decision of acceptance sampling, sample and sampling point were regarded as acceptable if the overall tracer concentration in an aggregate sample is more than 70% of the target dispensing rate and code-readable tracers from an individual grain source are included in the sample. These criteria were basically determined based on the lower limit of confidence intervals of tracer concentrations in the preliminary test and the observation from the prototype tracers printed with barcode.

2.5. Statistical analysis

All statistical analyses were performed using SAS software (2004). Randomness tests were performed to test the random distribution of tracers using the mean tracer contractions (Kendall & Smith, 1938). Simultaneous confidence interval construction with an individual binomial method and other methods proposed by Quesenberry and Hurst (1964), Goodman (1965), and Fitzpatrick and Scott (1987) were performed with an SAS macro using the PROC IML procedure (May & Johnson, 1997, 2000). The variance components of the tracer concentration and sample weight per probe trap in bins and trucks were estimated based on the nested repeated-measures ANOVA using Proc GLM of SAS. Three repeated samplings at each different location within a container were entered as a repeated measure.

3. Results and discussion

3.1. Overall assessment of the sampling process

The mean tracer concentrations observed in bins and trucks in the study were not much different from the target dispensing rate of three tracers per kg of wheat. The tracers were not highly concentrated in

Table 2
Nested repeated-measures ANOVA results on tracer concentration at different sampling points in the downscaled traceability system[a].

Source	Binomial probability distribution of tracers						Multinomial probability distribution of tracers					
	Elevator bin			Truck			Truck			Elevator bin		
	d.f	MS	F	d.f.	MS	F	d.f.	MS	F	d.f.	MS	F
Container[b]	4	5.4	0.6	6	27.5	**4.6**	5	14.2	2.3	1	42.0	5.8
Loc (container)[c]	20	11.3	1.2	28	6.4	1.1	36	9.0	1.4	8	7.1	0.5
Error	50	9.3		104			65			29		
Within subjects												
Rep[d]	1	1.6×10^7	**49,978**	1	2.1×10^7	**36,380**	6	3.0×10^6	**30,788**	6	38.5	**24.2**
Rep × container	4	2,341	**7.3**	6	7,566	**12.8**	30	3 923	**40.5**	6	6.6	**4.2**
Rep × Loc (container)	20	332	1.0	28	1,257	**2.1**	216	108	1.1	48	1.4	0.9
Error	50	320		70	591		144	97		120	1.6	

[a] The nested repeated-measures ANOVA was not performed for a large outdoor bin due to no nested data structure. Bold prints indicate statistical significance at $P < 0.01$.
[b] Either bin or truck.
[c] Loc = sampling location within a container.
[d] Rep = repeated samplings at each sampling location within a container.

Table 3
Mean tracer concentration (tracer/kg wheat) at different locations within a bin at dispensing rate of three tracers per kg of wheat[a].

Location	Bin (tracers/kg wheat)				
	# 1	# 2	# 3	# 4	# 5
1	3.39	4.50	2.55	2.02	5.93
2	1.43	0.51	4.14	5.46	5.79
3	4.85	2.01	1.62	0.53	2.62
4	1.98	4.03	7.20	0.51	1.59
5	3.97	2.49	1.05	2.11	2.48
Range	1.43–4.85	0.51–4.50	1.05–7.20	0.51–5.46	1.59–5.93

[a] Theoretically estimated 95% confidence intervals with binomial statistical methods as follows: 1.02–8.82 tracers/kg wheat (Wilson, 1927), −0.39–6.40 tracers/kg wheat (Wald, 1943), and 0.57–9.27 tracers/kg wheat (Agresti & Coull, 1998).

a particular area within a container, implying that tracers had not seriously segregated during grain handling and transporting. The estimated confidence intervals for tracer samples were largely influenced by the tracer concentration, sample size, and the statistical method used at the same significance level. A confidence interval is particularly important for the proposed traceability system to ensure whether sufficient samples have been taken to collect an adequate number of tracers (Hagstrum, Flinn, & Fargo, 1995). A theoretically computed sample size (kg) was empirically proven to be sufficient in collecting an appropriate number of tracers for the identification of grain sources. The results from actual sampling suggest that sample sizes (kg) could be lower than the calculated estimates by reducing the degree of precision in computing a sample size. However, care should be taken in defining precision to avoid the risk of missing tracers, particularly for a large number of grain sources. The tracer sampling showed different features as compared to typical grain sampling for highly skewed objectives or analytes (e.g., biotech grains and mycotoxin contaminated grains). During tracer sampling, the variation of tracer concentration in a representative sample might not be as critical as that in other grain sampling since even a few tracers can provide sufficient information on a specific grain lot. The designed sampling process was believed to provide a more representative measure of tracer concentration and distribution and thus increase sampling accuracy in identifying grain sources.

3.2. *Binomial probability distribution of tracers*

3.2.1. *Elevator bin sampling*

The first binomial probability sampling test was carried out in a bin, instead of in a transport moving from a field (Hirai, Schrock, Oard, & Herrman, 2006) since a tracer delivery could be efficiently monitored and controlled from the elevator under conditions of this study. This sampling test aimed to simulate bin sampling and its relevant activity in the commercial grain supply chain after a grain movement from a harvest field to a farm storage facility or to a country elevator. The sampling results indicated that tracers were not quite evenly distributed in grains within a bin. Consequently, the standard deviation and variation of tracer concentrations in an aggregate sample were large relative to the overall mean tracer concentration. However, the statistical results of randomness tests didn't show clear evidence of non-randomness and any recognizable patterns of average tracer concentrations on bin sampling and later truck sampling at $\alpha = 0.05$, indicating no significant segregation of tracers in a particular area within a container. The nested repeated-measures ANOVA results revealed that the tracer concentration was not significantly different from bin to bin and from location to location, but rather differed among locations within an individual bin and among repeated samplings at each location (Table 2). The whole model of tracer concentration built with bin and location nested within bin was not significant ($P = 0.4021$), explaining only 34% of the variance. The major portion of the variance in tracer concentration was explained by large variation in repeated sampling at sampling location within a bin. There was significant variation in sample weight per probe trap among the bins ($P < 0.01$). This variation indicates that the sample weight per probe trap was not very consistent from bin to bin. The variation in sample weight per probe trap is believed to be partially associated with the variation of the tracer concentration among the bins. This implies that the use of different sampling probes may in part alter the variance structure of the tracer concentration (Park, Whitaker, Giesbrecht, & Njapau., 2000).

With two exceptions, mean concentrations of the collected tracers at each sampling location within a bin fell within the confidence intervals estimated by the three binomial probability statistics (Table 3): the Wald interval, the Wilson interval, and the Agresti & Coull interval. Of these methods, the Agresti and Coull binomial interval at a 95% significance level covered most of the mean tracer concentrations found in a bin. Therefore, it is suggested the Agresti and Coull probability statistics should be used to estimate a confidence interval and the sample size (kg) for the binomial distribution of tracers. On the other hand, it is not anticipated that there are significant differences in the estimates of binomial proportions for the Agresti and Coull method and the other two methods.

3.2.2. *Truck sampling*

Grain and tracers in a truckload moving from the first sampling point to the large outdoor bin represent a binomial distribution. However, a difference existed between bin sampling and truck sampling although the overall mean tracer concentration between two locations slightly differed (0.26 tracer/kg of wheat). The nested repeated-measures ANOVA results showed a significant difference in the tracer concentrations among trucks and among repeated samplings ($P < 0.01$) (Table 2). However, the tracer concentration was not significantly different among locations within truck. The variation among locations within truck was smaller than that in the bin sampling. As observed for the bin sampling, the sample weight per probe trap appeared to be inconsistent from truck to truck

Table 4
Mean tracer concentration (tracer/kg wheat) at different locations within a truck at dispensing rate of three tracers per kg of wheat[a].

Location	Truck (tracers/kg wheat)						
	No 1	No 2	No 3	No 4	No 5	No 6	No 7
1	6.29	1.51	5.41	3.47	2.15	2.73	2.38
2	2.99	3.51	4.43	3.42	1.97	3.25	3.10
3	5.42	0.50	2.99	1.99	2.52	1.64	3.18
4	4.14	2.52	3.43	6.91	0.50	2.11	1.55
5	3.29	1.53	8.81	6.45	1.50	3.68	2.89
Range	2.99–6.29	0.50–3.51	2.99–8.81	1.99–6.91	0.50–2.52	1.64–3.68	1.55–3.18

[a] Estimated 95% confidence intervals are the same as those used for comparing tracer concentrations in a bin sampling: 1.02–8.82 tracers/kg wheat (Wilson, 1927), −0.39–6.40 tracers/kg wheat (Wald, 1943), and 0.57–9.27 tracers/kg wheat (Agresti & Coull, 1998).

($P < 0.01$), conceivably contributing to the variation in the tracer concentrations. Likewise, repeated sampling at sampling location within a truck was the largest source of variation in the truck sampling, explaining most of the variance in the tracer concentration (Table 2).

Table 4 shows that the lower limits of confidence intervals of the tracer concentrations within an individual truck are generally higher than those observed in bin sampling, indicating that it may be more feasible to collect an adequate number of tracers during truck sampling. The range of tracer concentrations in a truck was 0.50–8.81/kg of wheat, which roughly fell within the confidence intervals estimated by the three methods. As the results of the bin sampling showed, the ranges of tracer concentrations presented in Table 4 minimally exceeded the Agresti and Coull interval among the three methods. From the above observations, the truck sampling scheme for binomial populations appeared more feasibly to be correctly performed and implemented in the proposed traceability system.

The sample size (kg) for the binomial populations of tracers in a bin and a truck can be determined by properly converting the confidence interval formulas. The problem with estimated sample sizes (kg) in bin and truck sampling is the sampling location where the sample is not sufficient for capturing adequate tracers, for example, sampling locations with a mean tracer concentration of less than one tracer per kg of wheat in Tables 3 and 4. The way to solve this problem is to use more conservative binomial probability statistics, such as the Wald method and the Agresti and Coull method. Since mean tracer concentrations were calculated based on one kg of wheat, it is expected that even a small increase in sample size (e.g., 1–2 kg) is sufficient to collect a safe number of tracers at the present dispensing rate. Increased sample size (kg) can also effectively reduce the variability associated with sampling (Whitaker, 2006). The difference in sample sizes (kg) among statistical methods decreased with increasing target dispensing rate (data not shown). These observations imply that the choice of the right binomial probability method can select an appropriate sample size that ultimately optimizes sampling efficiency and cost in the actual traceability system.

3.3. *Multinomial probability distribution of tracers*

3.3.1. *Large outdoor bin sampling*

Grain and tracers transferred to the large outdoor bin from the five elevator bins initiated the mixing and commingling of grain. To identify all five grain sources by sampling, confidence intervals and sample size (kg)were estimated by using multinomial probability statistics suggested by Fitzpatrick and Scott (1987), Goodman (1965), Quesenberry and Hurst (1964), and a one-at-a-time binomial interval method. Table 5 presents the lower and upper limits of a 95% confidence interval for the number of tracers in the five bins at a concentration level of three tracers per kg of wheat. The Quesenberry and Hurst interval and the Goodman interval produced comparable and reasonable lower and upper limits for 95% confidence intervals, while the Fitzpatrick and Scott interval and the individual binomial interval were too conservative and unrealistic. The results in Table 5 indicate the importance of choosing appropriate statistical methods for the estimation of confidence intervals and the sampling design process.

Because the outdoor bin is relatively deep for the length of a sampling probe, the probe was only able to sample the upper grain surface. Furthermore, the grain transferred to the outdoor bin would typically be vertically stratified although the peak was leveled before sampling. As a result, probe samples from this bin failed to collect all five colored tracers (Table 6) even though sampling was undertaken in a rigorous fashion over the entire grain surface

Table 5
Construction of 95% simultaneous confidence intervals for multinomial distribution of tracers using different multinomial statistical methods, which allows one to estimate an approximate sample size (kg) to identify all of the five grain sources in an outdoor bin[a]

Sample size (kg)	Quesenberry and Hurst interval		Goodman interval		Fitzpatrick and Scott interval		Individual binomial interval	
	Lower limit ($\hat{p}_L$)	Upper limit ($\hat{p}_U$)	Lower limit ($\hat{p}_L$)	Upper limit ($\hat{p}_U$)	Lower limit ($\hat{p}_L$)	Upper limit ($\hat{p}_U$)	Lower limit ($\hat{p}_L$)	Upper limit ($\hat{p}_U$)
10	4.71×10^{-5} (16)[b]	1.56×10^{-4} (55)	5.32×10^{-5} (19)	1.38×10^{-4} (48)	-2.14×10^{-3} (0)	2.32×10^{-3} (810)	-1.82×10^{-3} (0)	1.99×10^{-3} (698)
20	5.60×10^{-5} (39)	1.31×10^{-4} (92)	6.11×10^{-5} (43)	1.20×10^{-4} (84)	-1.49×10^{-3} (0)	1.66×10^{-3} (1,164)	-1.26×10^{-3} (0)	1.43×10^{-3} (1,004)
30	6.05×10^{-5} (63)	1.22×10^{-4} (128)	6.50×10^{-5} (68)	1.13×10^{-4} (119)	-1.20×10^{-3} (0)	1.37×10^{-3} (1,442)	-1.02×10^{-3} (0)	1.19×10^{-3} (1,246)
40	6.33×10^{-5} (89)	1.16×10^{-4} (162)	6.74×10^{-5} (94)	1.09×10^{-4} (153)	-1.03×10^{-3} (0)	1.20×10^{-3} (1,681)	-8.68×10^{-4} (0)	1.04×10^{-3} (1,455)
50	6.54×10^{-5} (114)	1.12×10^{-4} (197)	6.91×10^{-5} (121)	1.06×10^{-4} (186)	-9.11×10^{-3} (0)	1.08×10^{-3} (1,895)	-7.67×10^{-4} (0)	9.39×10^{-4} (1,643)
100	7.08×10^{-5} (248)	1.04×10^{-4} (363)	7.36×10^{-5} (258)	9.98×10^{-5} (349)	-6.19×10^{-4} (0)	7.91×10^{-4} (2,768)	-5.18×10^{-4} (0)	$6.89x 10^{-4}$ (2,411)

[a] For the construction of a simultaneous confidence interval of tracers, it was assumed that the tracers are uniformly distributed in bulk grains and the tracer proportion ($\hat{p}_i$) at a concentration of three tracers per kg of wheat is identical for all different colored tracers. Confidence intervals were constructed using statistical methods proposed by Quesenberry and Hurst (1964), Goodman (1965), Fitzpatrick and Scott (1987), and the individual binomial method.

[b] Expected number of tracers in an aggregate sample.Negative lower limits resulting from impracticability of the statistical methods were exclusively converted to the zero tracer number.

Table 6
Overall mean tracer concentrations of different grain sources within a container at the latter three sampling points.

	Mean tracer concentrations of grain sources (tracers/kg of wheat)					
	Bin #1	Bin #2	Bin #3	Bin #4	Bin #5	Sum
Large outdoor bin	0.91	2.16	0.59	0.00	0.18	3.84
Truck	1.11	0.53	0.57	0.59	0.70	3.50
Two elevator bins	0.49	0.60	0.78	0.41	1.22	3.50

of the bin as indicated by the radial pattern in Fig. 4. Particularly, in case of the need for identifying many grain sources, such stratified and segregated grains at the sampling point should make it more difficult to collect all of the target tracers with great certainty.

The observed tracer concentration was not significantly different among the sampling locations without considering tracer color (grain source) (Fig. 4). The ANOVA test for the tracer concentration of each colored tracers also showed no significant difference among sampling locations, except for Bin 1's source ($P = 0.0377$). However, there was a large difference in tracer concentrations among different colored tracers in an aggregate sample (Table 6). Apparently, the sample weight per probe trap was not very consistent from the inner to the outer locations within the bin; lesser weights were observed at the outer locations.

The determination of an appropriate sample size for representing multinomial distributions of tracers and grains is not a simple task because it is closely associated with a fundamental problem in constructing a confidence interval for multinomial proportions (Angers, 1974 & 1984, Thompson, 1987; Tortora, 1978). Of the statistical methods employed for estimating a sample size, the Angers method and the Tortora method produced a practical sample size (kg) and were equally simple and easy to use with an Excel spreadsheet. Contrarily, the Thompson (1987) method with the simplest computational procedure was too conservative and required a much larger sample size (kg) (data not shown). The large difference in the estimated sample sizes (kg) among the statistical methods can be partially attributed to a much larger proportion of grain as compared to that of tracers. Table 7 shows a proportional relationship between sample size (kg) and a number of grain sources in bulk grain at two target dispensing rates. According to this computation, a dispensing rate of one tracer per kg of wheat requires 2–3 times the sample size (kg) as compared to three tracers per kg of wheat. This sample size (kg) should change further with the magnitude of variation in the tracer concentration for actual sampling.

The estimated sample size for identifying all five grain sources was 15.3 kg according to the Angers (1984) and Tortora (1978) methods. On the average, this sample size is expected to give approximately 45.9 tracers (15.3 kg $\times$ 3 tracers per kg of wheat). However, in actual sampling, the calculation based on mean tracer concentrations in Table 6 at the given sample size of 15.3 kg indicates that the expected tracers from each bin would be 14 tracers from Bin 1, 33 tracers from Bin 2, nine tracers from Bin 3, three tracers from Bin 5, and no tracers from Bin 4, resulting in a total of 59 tracers. To collect 59 tracers under the same sampling conditions, a sample size of about 19.7 kg (59 tracers/3 tracers per kg of wheat) would be required. The absence and very small amount of tracers from particular bin sources indicate that the sample and sampling point are not quite acceptable according to the predetermined criteria for acceptance sampling. The above observation also suggests that the actual sample size can be lowered from the estimated sample size. However, if the sample size is too small, the information contained in coded tracers would be lost. The required smaller sample size (kg) for the actual sampling may be attributed to a more homogeneous distribution of tracers when compared to sampling for skewed target objectives or analytes, which generally require larger sample sizes. From the sampling results, it is noteworthy to mention that even a small number of tracers obtained from a single grain source (e.g., Bin 5) might be sufficient to retrieve adequate data if the tracers are not damaged and the bar codes are readable by a scanner. For the estimation of the sample size (kg), a predetermined value close to a proportion of each colored tracer in bulk grain was used as the acceptable width (anticipated precision) of intervals for the sample size formulas. Since the use of this precision value resulted in a slightly larger sample size (larger number of tracers) than what is necessary for actual sampling, it is advised that the sample size (kg) be estimated with a decreased degree of precision to take into consideration of the difference in the number of tracers between the estimated and actual values.

3.3.2. Truck sampling

Unlike sampling in the large outdoor bin, multiple samples taken from six truckload transfers from the large outdoor bin to the two elevator bins included all five colored tracers and thereby enabled the identification of all of the grain sources (Table 6). The nested repeated-measures ANOVA results show that the tracer concentration, without considering the tracer color, was not significantly different among trucks and locations nested within truck. However, some colored tracers were highly concentrated in a specific truck and distributed at a specific location within each truckload. This was, to some extent, anticipated because grains from the five elevator bins were sequentially delivered and stored in the outdoor bin. Because the different grain sources did not have a chance to mix with each other within the large outdoor bin, individual truckloads hauling about 300 bu of the vertically-stratified grain should contain a larger portion of a single grain source.

The aggregate sample for six trucks with a combined sample size of 15.3 kg included 17 tracers for Bin 1, eight tracers for Bin 2,

Table 7
A computation of required sample sizes (kg) to ensure collection of a representative sample of grains at target dispensing rates of one and three tracers per kg of wheat.[a]

Number of grain sources (bins)[b]	Tracer concentration (tracer/kg of wheat)[b]	Tracer proportion ($\widehat{p}_i$)[c]	Required sample size (kg)[d]		Number of truck units[e]
			Angers method	Tortora method	
5	1	5.71×10^{-6}	36.8	36.8	7
	3	1.71×10^{-5}	12.3	12.3	
10	1	2.86×10^{-6}	73.5	73.5	14
	3	8.57×10^{-6}	24.5	29.1	
20	1	1.43×10^{-6}	147.1	147.0	27
	3	4.29×10^{-6}	49.0	67.5	

[a] Some assumptions made for a sample size computation are same as those made in Table 4.
[b] Each bin capacity was assumed as 10,935 kg used for the probability sampling test.
[c] This represents the proportion of tracers from a single bin in bulk grains.
[d] Required sample sizes (kg) were computed using formulas proposed by Angers (1974, 1984) and Tortora (1978).
[e] The number of truck units to haul grains was calculated based on the truck capacity of 8100 kg (300 bu).

nine tracers for Bin 3, nine tracers for Bin 4, and 11 tracers for Bin 5 for actual sampling, resulting in a total of 54 tracers. The mean concentration and distribution of tracers in the aggregate sample were acceptable according to acceptance sampling criteria. The overall tracer concentration without considering the tracer color was also slightly lower in a truckload (3.50 per kg wheat) than in the outdoor bin (3.84 per kg of wheat). However, different colored tracers in an aggregate sample were more evenly distributed as compared to the outdoor bin sampling (Table 6). Again, the sample size (kg) determined by the probability statistics appeared to be sufficient to capture an adequate number of tracers, and it may even be slightly overestimated. To make an estimate that is closer to a practical sample size (kg) and an optimized number of tracers for grain source identification, the degree of precision may be lowered in the actual traceability system. These observations and inferences demonstrate better efficiency and convenience of truck sampling as compared to bin sampling for the multinomial proportion of tracers in terms of the homogeneity of the tracer distribution and practical sampling effectiveness.

3.3.3. Two elevator bin sampling

The overall mean tracer concentration from the two elevator bins was 3.5 tracers per kg of wheat, which is similar to that observed in the previous truck samplings (Table 6). However, the mean tracer concentration was somewhat different between Bin A (4.7 tracers/kg of wheat) and Bin B (2.3 tracers/kg of wheat). Interestingly, all five colored tracers were identified in Bin A, although the relatively small sample size (kg) was taken on the top grain surface at only five locations. This indicates that tracers in the grain had become commingled during the sequential grain handling operations and transportation. The nested repeated-measures ANOVA results show that the overall tracer concentration and individual colored tracer concentrations were not different among the sampling locations within bin (Table 2). Similar to other sampling points, the repeated sampling at sampling location within a bin were a greater source of variation in tracer concentration, but the variation was much smaller than at other sampling points. The estimated sample size of 15.3 kg for the multinomial proportion of tracers was sufficient for collecting all five colored tracers and identifying all of the grain sources. At this given sample size (kg), a total of 54 tracers could be drawn from both bins, including eight tracers from Bin 1, nine tracers from Bin 2, 12 tracers from Bin 3, six tracers from Bin 4, and 19 tracers from Bin 5.

The multinomial sampling results suggest that the increasing frequency of probing at more sampling locations presents a better strategy for capturing tracers than increasing the sampling size at fewer sampling locations, as also indicated in the sampling study of Hart and Schabenberger (1998). In the commercial grain supply chain, the internal movement of grain from bin to bin, which often occurs due to storage space and grain quality management, provides more chances for the mixing of different of grain lots and thereby provides better tracer identification of grain sources. In the event that multinomial proportions of tracers are unknown, pre-sampling may be conducted to determine the actual multinomial distribution of grain. If computed sample sizes (kg) based on pre-sampling information are less than the corrected sample sizes, extra samples may be taken, thus avoiding wasteful sampling (Angers, 1984).

In conclusion, binomial and multinomial probability statistics explained tracer distribution in bulk grain at different sampling points within this study. The suggested statistical sampling strategies and principles can be readily used for any sampling point and condition in a typical grain supply chain by the slight modification of sampling procedures and spreadsheet formula parameters that were developed in this study. Accordingly, knowledge of the applied statistics and parameters of the probability distribution of tracers in grain lots should be well understood. Some other important aspects, such as the number of samples, the selection of sampling probes, appropriate sampling locations, the variance structure of tracer concentration, sampling precision, and the homogeneity of tracer distribution should also be considered in designing a tracer sampling process to improve the sampling reliability.

Preliminary projected costs for a tracer-based system are projected as follows. The number of tracers used to mark 100,000 kg of grain would require 100,000 tracers (approx. 36 kg) and 300,000 tracers (108 kg) at the concentrations of one and three tracers per kg of wheat, respectively. At one tracer per kg of grain and a cost of $0.001 per tracer (0.1 cent), the tracer cost per bushel of wheat is approximately $0.027. If expanded to the entire U.S. grain crop of approximately 2.2 billion bushels, the annual cost of tracers would approach $59 million. Fixed costs for dispensing, extraction and tracer identification equipment plus management costs would add to this.

Applying probability theory to the tracer sampling design process would help system participants plan a more efficient and practical sampling strategy to reduce economic risks in grain marketing. While the proposed traceability system cannot prevent food safety or bio-security hazards, it would greatly facilitate trace-forward and trace-back activities to help identify the point where the hazard was introduced. The rapid identification of the source of contamination is essential in helping define the path of a bio-terror, bio-hazard, or food-borne disease outbreak and the subsequent recovery activities.

Acknowledgment

The authors acknowledge that this research was supported by the National Research Initiative of the USDA-CSREES Plant Biosecurity Program.

Appendix. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.foodcont.2010.12.016.

References

Agresti, A., & Coull, B. A. (1998). Approximate is better than “exact” for interval estimation of binomial proportions. *The American Statistician, 52*(2), 119–126.
Angers, C. (1974). A graphical method to evaluate sample size for the multinomial distribution. *Technometrics, 16*. 469–471.
Angers, C. (1984). Large sample sizes for the estimation of multinomial frequencies from simulation studies. *Simulation*, 175–178, (October).
Can-trace. (2003). *Agriculture and agri-food Canada*. Available from. http://www.can-trace.org.
Coker, R. D., Nagler, M. J., Blunden, G., Sharkey, A. J., Defize, P. R., Derksen, G. B., & Whitaker, T. B. (1995). Design of sampling plans for mycotoxins in foods and feeds. *Natural Toxins, 3*, 257–262.
ERS (Economic Research Service of the U.S. Department of Agriculture). (2001). *Knowing where it's going: Brining food to market in the age of genetically modified crops*. Available from. http://www.pewtrusts.org/uploadedFiles/wwwpewtrusts org/Summaries_-_reports_and_pubs/hhs_biotech_knowing.pdf.
Federal Register. (2004). Establishment and maintenance of records under the public health security and bioterrorism preparedness and response act of 2002. *Federal Register, 69*(Dec 9, 236), 71561–71655, Final Rule.
Fitzpatrick, S., & Scott, A. (1987). Quick simultaneous confidence intervals for multinomial properties. *Journal of the American Statistical Association, 82*(399), 875–878.
Golan, E., Krissoff, B., Kuchler, F., Calvin, L., Nelson, K., & Price, G. (2004). *Traceability in the U.S. food supply: Economic theory and industry studies*. Agricultural Economic Report No. 830. Washington, D.C.: USDA/Economic Research Service.
Goodman, L. A. (1965). On simultaneous confidence intervals for multinomial proportions. *Technometrics, 7*(2), 247–254.
Hagstrum, D., Flinn, P., & Fargo, S. (1995). *How to sample grain for insects. Stored product management*. Oklahoma State University. Cooperative Extension Service Circular E-91265–69.

Hart, L. P., & Schabenberger, O. (1998). Variability of vomitoxin in truckloads of wheat in a wheat scab epidemic year. *Plant Disease, 82*(6), 625–630.

Hellevang, K. J., Backer, L. F., & Maher, G. G. (1992). *Grain stream sampling and sampler construction*. North DakotaState University Extension Service. Bulletin AE-1044.

Herrman, T. J. (2002). *White paper on traceability in the U. S. grain and plant protein feed ingredient industries*. Available from. http://www.oznet.ksu.edu/grsiext/White%20Paper%20CVM.PDF.

Hirai, Y., Schrock, M. D., Oard, D. L., & Herrman, T. J. (2006). Delivery system of tracing caplets for wheat grain traceability. *Applied Engineering in Agriculture, 22*(5), 747–750.

ISO 22005. (2007). *Traceability in the feed and food chain – General principles and guidance for system design and development*. ISO Standard.

Kendall, M. G., & Smith, B. B. (1938). Randomness and random sampling numbers. *Journal of the Royal Statistical Society, 101*(1), 147–166.

Larsen, R. J., & Marx, M. L. (1985). Combinatorics. In R. J. Larsen, & M. L. Marx (Eds.), *An introduction to probability and its applications* (pp. 96–168). New Jersey: Prentic-Hall, Inc.

May, W. L., & Johnson, W. D. (1997). A SAS macro for constructing simultaneous confidence intervals for multinomial proportions. *Computer Methods and Programs in Biomedicine, 53*, 153–162.

May, W. L., & Johnson, W. D. (2000). Constructing two-sided simultaneous intervals for multinomial proportions for small counts in a large number of cells. *Journal of Statistical Software, 5*(6), 1–24.

Meyer, S. (2004). *Grain transportation and marketing channels*. Available from. http://www.fapri.missouri.edu/outreach/publications/2004/FAPRI_UMC_Briefing_Paper_04_04.pdf.

Official Journal of the European Communities. (2002). *Regulation (EC) No. 178/2002 of the European Parliament and the Council of 28 January 2002*.

Park, D. L., Whitaker, T. B., Giesbrecht, F. G., & Njapau, H. (2000). Performance of three pneumatic probe samplers and four analytical methods used to estimate aflatoxins in bulk cottonseed. *The Journal of AOAC International, 83*(5), 1247–1251.

Quesenberry, C. P., & Hurst, D. C. (1964). Large sample simultaneous confidence intervals for multinomial proportions. *Technometrics, 6*(2), 191–195.

SAS. (2004).) *SAS user's guide: Statistics (version 9.2*. Cary, NC. U.S.A: SAS Institute Inc.

Sims, F. (2005). A commercial approach to managing the supply chain. *Proceedings of Institute of Food Technologist's First Annual Food Protection and Defense Research Conference*. Atlanta, Georgia.

Thakur, M., & Hurburgh, C. R. (2009). Framework for implementing traceability system in the bulk grain supply chain. *Journal of Food Engineering, 95*, 615–626.

Thompson, S. K. (1987). Sample size for estimating multinomial proportions. *The American Statistician, 41*(1), 42–46.

Tortora, R. D. (1978). A note on sample size estimation for multinomial populations. *The American Statistician, 32*(3), 100–102.

U.S. Congress. (2002). *Public health security and bioterrorism preparedness and response act of 2002*. Title III, subtitles A. Washington, D.C., June 2002. Available from. http://www.fda.gov/oc/bioterrorism/bioact.html.

USDA (U.S. Department of Agriculture). (2005). *Global traceability and labeling requirements for agricultural biotechnology-derived products: Impacts and implications for the United States*. Washington, D.C., May 2005. Available from. http://www.usda.gov/wps/portal/usdahome?contentidonly=true&contentid=event_15.xml.

Vachal, K., & Tolliver, D. (2001). *Regional elevator survey: Grain transportation operation and industry trends for Great Plains elevators*. Available from. http://ntlsearch.bts.gov/tris/record/ntl/12368.html.

Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society, 54*(3), 426–482.

Whitaker, T. B. (2006). Sampling food for mycotoxins. *Food Additives and Contaminants, 23*(1), 50–61.

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association, 22*(158), 209–212.