# The compartment bag test (CBT) for enumerating fecal indicator bacteria: basis for design and interpretation of results

Andrew D. Gronewold [a,b,*], Mark D. Sobsey [c], Ku McMahan [d]

[a] *NOAA, Great Lakes Environmental Research Laboratory, Ann Arbor, Michigan, USA*

[b] *Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, Michigan, USA*

[c] *Department of Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina Chapel Hill, Chapel Hill, North Carolina, USA*

[d] *United States Agency for International Development (USAID), Washington, D.C., USA*

**Abstract**

For the past several years, the compartment bag test (CBT) has been employed in water quality monitoring and public health protection around the world. To date, however, the statistical basis for the design and recommended procedures for enumerating fecal indicator bacteria (FIB) concentrations from CBT results have not been formally documented. Here, we provide that documentation following protocols for communicating the evolution of similar water quality testing procedures. We begin with an overview of the statistical theory behind the CBT, followed by a description of how that theory was applied to determine an optimal CBT design. We then provide recommendations for interpreting CBT results, including procedures for estimating quantiles of the FIB concentration probability distribution, and the confidence of compliance with recognized water quality guidelines. We synthesize these values in custom user-oriented 'look-up' tables similar to those developed for other FIB water quality testing methods. Modified versions of our tables are currently distributed commercially as part of the CBT testing kit.

*Key words:*
Compartment bag test; water quality; drinking water; human health; statistical methods

## 1 Introduction

Ensuring readily-available high quality drinking water is fundamental to human health and has important connections to socioeconomic status, commercial and industrial growth, and overall quality of life (Mekonnen and Hoekstra, 2016). The challenge of providing that ensurance is met in different ways around the world; in some communities, drinking water supplies are assumed protected if they are adequately separated from wastewater and other sources of contamination (George, 2008). In others, routine water quality testing is used to ensure compliance with recognized standards (Gleick, 1998; Novotny, 2003). Testing kits that support these assessments often require a skilled technician to collect, analyze, and interpret results, as well as microbiological laboratory facilities. In regions of the world without these resources and where the time from water withdrawal (from its source) to consumption is short, alternative testing procedures are needed.

To address this gap in global water quality protection, researchers at the University of North Carolina Chapel Hill and Duke University developed a simple kit for enumerating FIB concentrations that is portable, relatively inexpensive, and provides easy-to-interpret results (Stauber et al., 2014). This kit, commonly referred to as the compartment bag test (or CBT), is currently manufactured and distributed by Aquagenx, LLC and has been tested and used in communities around the world (Murcott et al., 2015; Weiss et al., 2016). To date, however, the statistical basis for the design and recommended interpretation of results from the CBT have not been formally documented.

Here, following documentation for the development of similar water quality testing kits (McCrady, 1915; de Man, 1977; Tillett and Coleman, 1985; Haas, 1989; McBride et al., 2003), we begin with an overview of the statistical the-

* Corresponding author. Tel.: +1 734 741 2444, Fax: +1 734 741 2055
*Email address:* drew.gronewold@noaa.gov (Andrew D. Gronewold).

ory behind the CBT, followed by examples of how that theory was applied to determine an optimal CBT design. We then provide recommendations for interpreting CBT results, including procedures for estimating quantiles of the FIB concentration probability distribution, as well as procedures for calculating the confidence of compliance with World Health Organization (WHO) drinking water quality guidelines (McBride and Ellis, 2001; Borsuk et al., 2002; World Health Organization, 2004). We synthesize these values in custom user-oriented 'look-up' tables similar to those developed for other FIB testing kits (de Man, 1977). Finally, we explore the sensitivity of CBT results to departures from assumptions in the underlying statistical models, and from recommended protocols for sample collection and handling.

## 2 Experimental

### *2.1 Statistical basis for interpreting CBT results*

The CBT is a manufactured clear plastic multi-compartment bag into which 100 ml of a water sample is distributed (Stauber et al., 2014). Each compartment contains a growth substrate designed to detect groups of FIB (such as hydrogen sulphide producers), or specific bacteria such as *Escherichia coli* (EC), by turning a distinctive color (e.g. blue-green) indicating growth of "target" (e.g. FIB or EC) bacteria during an incubation period. The CBT will yield a pattern of 'positive' and 'negative' compartments from which a user can infer the FIB concentration of the original sample following the common assumption (Greenwood and Yule, 1917; Cochran, 1950; Woodward, 1957; El-Shaarawi et al., 1981; Hurley and Roscoe, 1983; de Man, 1983; Haas and Heller, 1988; Woomer et al., 1990; Briones and Reichardt, 1999) that, for a given sample, the number of target bacteria ($y_i$) in compartment $i$ ($i \in [1,m]$ and $m$ is the total number of compartments) with volume $v_i$ (assuming a well-mixed sample) is well-represented by a Poisson probability distribution

$y_i \sim \mathsf{Po}(\lambda_i = cv_i/100)$ with FIB concentration $c$ (in organisms per 100 ml), and mean and variance $\lambda_i$. The probability of a positive compartment of volume $v_i$ is 1-exp(-$cv_i$/100). The joint probability of any pattern of positive and negative compartments $\vec{x}$ (where the over-arrow superscript denotes a row vector, $x_i \in$[0,1] and $x$=1 indicates a positive compartment) is then expressed as the product of a series of $m$ independent Bernoulli trials:

$$f(\vec{x} \mid \vec{v}, c) \propto \prod_{i=1}^{m} \left(1 - e^{-cv_i/100}\right)^{x_i} \left(e^{-cv_i/100}\right)^{1-x_i} \tag{1}$$

Conventional interpretations of presence/absence test kits for FIB often focus on a deterministic solution to $c$ from equation 1. This value is commonly referred to as the "most probable number" (or MPN) and can be calculated as (Hurley and Roscoe, 1983; McBride, 2005; Gronewold and Wolpert, 2008):

$$\text{MPN} = \underset{c}{\text{argmax}} \left[ \prod_{i=1}^{m} \left(1 - e^{-cv_i/100}\right)^{x_i} \left(e^{-cv_i/100}\right)^{1-x_i} \right] \tag{2}$$

We implement this formulation using the `uniroot` function in the `R` statistical software package (R core team, 2014). Corresponding code is included in the Supplementary Information.

Multiple methods have been developed for expressing uncertainty in the MPN, however most do not explicitly acknowledge that the probability distribution of the MPN for a given pattern of positive and negative compartments is typically discrete and multi-modal, while the probability distribution of the FIB concentration is almost always unimodal and continuous (Klee, 1993; Gronewold and Wolpert, 2008). Therefore, in addition to reporting conventional MPN values, we propose two interpretations of CBT results that allow for a more robust understanding of the uncertainty in the FIB concentration and how that uncertainty affects the confidence of compliance with water quality guidelines (McBride and Ellis, 2001; Gronewold and Borsuk, 2009, 2010).

77 The first is based on calculating quantiles of the likelihood function of the FIB
78 concentration (equation 1, written as a function of $c$ for given $\vec{x}$ and $\vec{v}$), as
79 well as the probability that the FIB concentration exceeds 1, 10, 100, or 1000
80 organisms per 100 ml.

81 The second interpretation is based on a Bayesian analysis of CBT results
82 (Bernardo and Ramon, 1998; Press, 2003; Bolstad, 2004) where the posterior
83 probability distribution of the FIB concentration $c$ is proportional to the prod-
84 uct of the likelihood function (equation 1) and prior probability distribution
85 $\pi(c)$:

$$f(c \mid \vec{x}, \vec{v}) \propto \pi(c) f(\vec{x} \mid \vec{v}, c) \tag{3}$$

86 One advantage of this approach is that it allows for expression of *a priori*
87 assumptions about the potential range of the FIB concentration in a water
88 sample. Methods based on the likelihood function alone, in contrast, implicitly
89 assume *a priori* that FIB concentrations ranging from 0 to $\infty$ are equally
90 likely; an assumption analogous to a belief that gross contamination is just
91 as likely as a FIB concentration within a few orders of magnitude of (or even
92 well below) WHO water quality guidelines. This *a priori* belief is just one of
93 many a CBT user might have about water quality at a particular sampling
94 location (Press, 2003). Here, we present calculations based on a lognormal
95 prior $\pi(c) = \mathsf{LN}(\mu = 0, \sigma^2 = 100)$, with log-concentration mean $\mu$ and variance
96 $\sigma^2$, intended to represent an *a priori* belief that the FIB concentration is most
97 likely low, but that extreme FIB concentrations are possible. We view further
98 investigation of impacts of alternative priors on CBT results as an important
99 area for future research.

100 It is informative to note that previous studies have explored alternative prob-
101 ability models for interpreting multiple-compartment water quality analysis
102 results, including the negative binomial model and variations of the Poisson

model that account for thinning and dispersion (Christian and Pipes, 1983; El-Shaarawi et al., 1981; Messner and Wolpert, 2002; Crainiceanu et al., 2003). Recent research, however (see Gronewold et al., 2008; Wu et al., 2014), indicates that only extreme and persistent violations of the Poisson probability model would justify application of an alternative probability model.

Finally, following equation 1, we calculate the relative likelihood of each possible combination of positive and negative compartments. Results of this calculation provide an indication of CBT outcomes that are most likely, and those that (because they are extremely unlikely) might indicate contamination or thinning of individual compartments and would therefore warrant additional testing and verification.

### *2.2 Design criteria*

The number and volume of compartments of the CBT is based on consideration of a range of criteria including ease of manufacturing, minimization of potential user error (such as unintentionally distributing more or less water into each CBT compartment than intended), and results that are readily translatable into health risk-based metrics. More specifically, the ideal CBT design yields a pattern of positive and negative compartments that are easy to translate into FIB concentrations with uncertainty bounds relevant to human health risks. For most applications of the CBT, we expect these risks will be assessed using FIB concentration numeric limits prescribed in WHO water quality guidelines. We assess compliance with this criteria by inferring FIB concentrations associated with each possible result (i.e. each combination of positive and negative compartments) of a particular CBT design, and then comparing these concentrations to established water quality criteria and standards.

To demonstrate our approach, we provide a comparison between two CBT

designs. The first (the design ultimately employed in practice) is a CBT with five compartments with volumes (in ml) $\overrightarrow{v} = \{56, 30, 10, 3, 1\}$. The second is a CBT with seven compartments with volumes $\overrightarrow{v} = \{37, 32, 16, 8, 4, 2, 1\}$. These design options evolved out of a qualitative consideration of the aforementioned criteria, as well as the constraints that the cumulative volume of all compartments equal 100ml, and that the compartment volumes span as broad a range as possible without multiple compartments of the same volume.

For each of the two test designs, we first calculated the full FIB concentration likelihood function for each possible CBT result, and then implemented our Bayesian interpretation by simulating samples from the posterior probability distribution of the FIB concentration (equation 3) for each possible CBT result using Markov chain Monte Carlo (MCMC) procedures in the software program WinBUGS (Lunn et al., 2000). We ran each MCMC chain until it reached convergence, indicated by a potential scale reduction factor $\hat{R}$ (Gelman et al., 2004) close to 1.0. WinBUGS code used to simulate the posterior probability distribution for $c$ for the $\overrightarrow{v} = \{56, 30, 10, 3, 1\}$ CBT design is included in the Supplementary Information. From the likelihood functions and posterior probability distributions, we calculate a series of quantiles, as well as the likelihood (or posterior probability) that the FIB concentration exceeds 1, 10, 100, or 1000 organisms per 100 ml.

### *2.3 Sensitivity analysis*

To better understand the sensitivity of CBT results to potential variations in user handling (including violations of the assumptions in our statistical models), we repeat the simulation described in the previous section for the 5-compartment CBT using hypothetical compartment volumes (in ml) of $\overrightarrow{v} = \{58.4, 30.5, 14.5, 2.5, 0.7\}$ and $\overrightarrow{v} = \{32.3, 33.5, 23.3, 4.9, 3.4\}$. These volume sequences were obtained from an informal (unpublished) study by one of the

authors at the University of North Carolina - Chapel Hill in which roughly twenty individuals with a range of CBT experience used the CBT, and the actual water sample volumes they distributed into each compartment were recorded. The two selected sequences represent, respectively, moderate and severe departures from the intended 5-compartment CBT design with compartment volumes $\overrightarrow{v} = \{56, 30, 10, 3, 1\}$.

## 3 Results and discussion

Of the 32 potential combinations of positive and negative compartments for the 5-compartment CBT, we find that there are appreciable differences in the relative likelihood of each outcome (see Table S1 in Supplementary Information). Some results (particularly those for which the 56ml compartment is positive) are quite likely while others are highly improbable. This is an important distinction because highly unlikely CBT outcomes might indicate one or more potential problems with sample handling and analysis (including thinning or contamination of a particular compartment) and warrant additional investigation. To underscore this point, and to simplify our discussion of alternative CBT interpretations, we hereafter focus on results from only the eight most likely outcomes of the CBT.

FIB concentration likelihood functions reflecting information content of individual CBT compartments (top five rows figure 1), and of each combination of positive and negative compartments for the eight most likely CBT results (bottom row figure 1), provide insight into origins of uncertainty in CBT-based water quality assessments (see also table 1). For example, a CBT result with a pattern of positive (1) and negative (0) compartments (with volumes 56, 30, 10, 3, and 1ml) of $\overrightarrow{x} = \{1, 1, 0, 1, 0\}$ has an MPN of 9.6 (organisms per 100 ml) with moderate certainty in the FIB concentration. A CBT result for which the pattern of positive and negative compartments is $\overrightarrow{x} = \{1, 1, 1, 1, 0\}$

has a higher MPN (48.3) and more uncertainty in the FIB concentration because of the difference in the information content of the 10ml compartment. A positive 10ml compartment (by itself) indicates that the FIB concentration is almost certainly above roughly 40 organisms per 100 ml, while a negative 10ml compartment indicates that the concentration is almost certainly below 40 organisms per 100ml. The contrast between the information in these two results underscores not only the relative value of keeping the CBT simple (by minimizing the number of compartments, for example) and easy to implement, but also the potential sensitivity of CBT outcomes to variations in sample handling.

A Bayesian interpretation of results from the 5-compartment CBT with $\overrightarrow{v} = \{56, 30, 10, 3, 1\}$ (figure 2 and table 2) indicates how explicit quantification of *a priori* beliefs about the FIB concentration in a sample can propagate into different perceptions of human health risk (figure 2) when compared to interpretations based on the likelihood function alone, particularly for CBT outcomes with an intrinsically broad likelihood function (e.g. a result of $\overrightarrow{x} = \{1, 1, 1, 1, 0\}$). In areas where there is a long history of high quality drinking water, for example, a prior probability distribution reflecting a strong belief in a relatively low FIB concentration may be helpful in guiding water use management decisions when there is insufficient information content in the likelihood function alone.

We also find that the 5-compartment CBT design (tables 1 and 2) provides a robust basis for distinguishing samples based on compliance with WHO water quality guidelines, particularly when compared to our alternative design with seven compartments (see table S2 in Supplementary Information). For nearly all of the most likely results of the 5-compartment CBT, we can make a relatively confident statement about the range of the sample FIB concentration, and about compliance with each numeric limit in the WHO guidelines. This statement may depend, as we have shown, on whether a likelihood or Bayesian

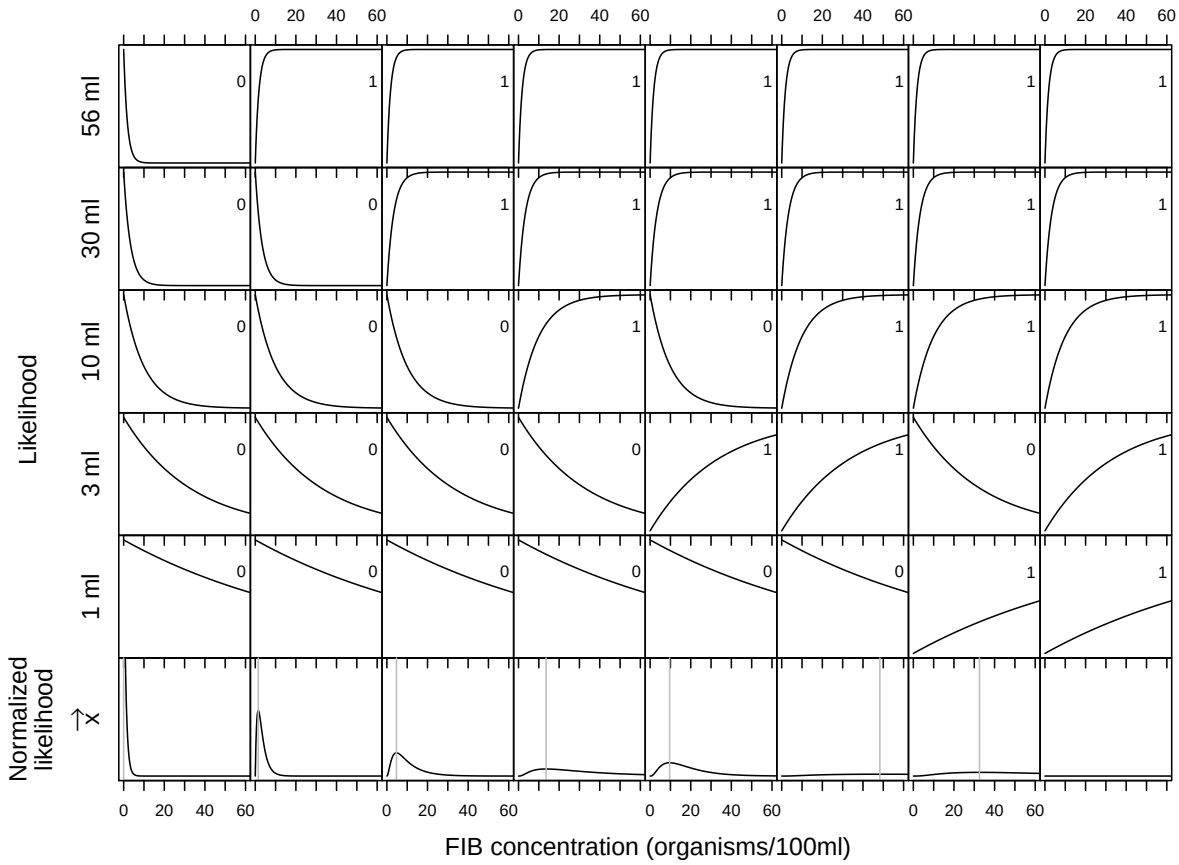



Fig. 1. Likelihood functions (first five rows) for individual positive ('1' in upper-right corner of panel) or negative ('0' in upper-right corner) compartments of the 5-compartment CBT, and normalized likelihood functions (bottom row) for the eight most likely outcomes of the CBT. Likelihood functions in the bottom row reflect the combined results of the positive and negative compartments from the five panels above (from the same column). The bottom left-most panel, for example, is the normalized likelihood function for the FIB concentration from a CBT result with all compartments negative. Vertical grey lines in the panels of the bottom row indicate the MPN (note that the MPN is undefined when all compartments are positive).

interpretation is used. In either case, a probabilistic interpretation enhances information from conventional MPN values alone; water quality experts are often comfortable with MPN values, but not with quantifying associated uncertainties when the MPN is derived from a novel and unconventional testing kit such as the CBT.

Our assessment of the potential impacts of user error (table 3 and Supplementary Information) suggests that the 5-compartment CBT test is relatively robust to both moderate and severe errors. More specifically, we find that

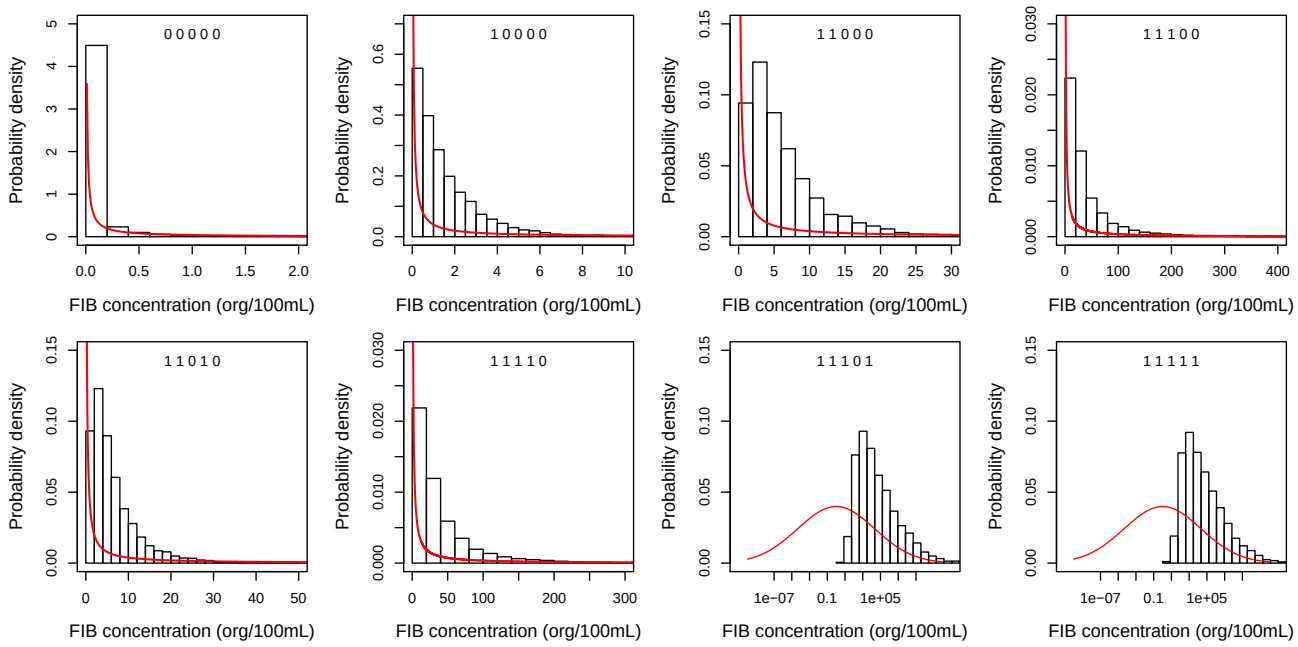



Fig. 2. Bayesian interpretation of CBT results including FIB concentration prior probability distribution (red lines) and histograms of simulated samples from the FIB concentration posterior probability distribution for the eight most likely results from the 5-compartment (volumes 56, 30, 10, 3, 1 ml) CBT. Values of 1 and 0 across the top of each panel correspond to each pattern of positive (1) and negative (0) compartments of volumes 56, 30, 10, 3, 1 ml, respectively. The x-axis of the two right-most panels in the bottom row is plotted on a logarithmic scale for clarity.

moderate handling errors would not have changed the perceived probability of violating the WHO water quality guideline of 100 organisms per 100 ml (a value indicating 'very high risk' water). Furthermore, we find that severe errors, while leading to a slightly lower perceived probability of violating the WHO water quality guideline of 100 organisms per 100 ml, would also have been very unlikely to lead to a different perception of risk than what would have been inferred had there been no error.

Finally, we acknowledge that users of the CBT have inquired about the uncertainty in CBT results relative to uncertainties in more conventional water quality testing tools, including (for example) membrane filtration (MF) tests (Dufour and Cabelli, 1975; Dufour et al., 1981; El-Shaarawi et al., 1981). A comparison between the 95% likelihood intervals from our analysis of the CBT (table 1) and 95% likelihood intervals from MF tests with colony-forming unit

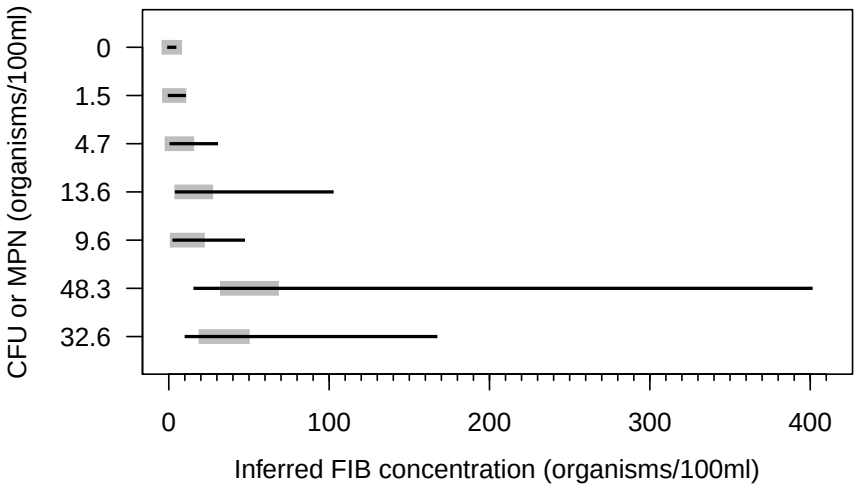



Fig. 3. FIB concentration 95% likelihood intervals based on seven of the eight most likely results of the 5-compartment CBT (intervals for a CBT result with all compartments positive are not shown because the likelihood function is continuously increasing and the MPN is undefined). Thin black segments represent likelihood intervals derived from the CBT. Thick grey segments represent likelihood intervals derived from conventional membrane filtration (MF) analyses with CFU values that correspond to MPN values from the CBT. The top-most pair of segments, for example, includes FIB concentration 95% likelihood intervals from (thin black segment) a CBT result with an MPN of 0, and (thick grey segment) an MF result with a CFU of 0.

(CFU) values matching MPN values from the CBT (Gronewold and Wolpert, 2008) indicates that (figure 3), for very low (i.e. less than 5 organisms per 100ml) FIB concentrations, the confidence intervals are quite similar and that the differences are more extreme for FIB concentration close to and above 10 (organisms per 100ml). A Bayesian interpretation of CBT results (table 2) could affect the range of these intervals and might in fact be desirable should water quality management officials (and other CBT users) find that the likelihood-based intervals do not provide enough informative at higher concentrations. We suggest investigation of impacts of alternative prior distributions on inferred FIB concentration uncertainty and compliance with WHO water quality guidelines as a high priority for future research.

## 4 Acknowledgments

The authors thank Craig Stow for comments on early versions of this manuscript, as well as Song Qian and James Crooks for helpful discussions on Bayesian inference. This is NOAA-GLERL contribution number XXXX.

## References

Bernardo, J. M., Ramon, J. M., 1998. An introduction to Bayesian reference analysis: inference on the ratio of multinomial parameters 47 (1), 101–135.

Bolstad, W. M., 2004. Introduction to Bayesian Statistics. Wiley-Interscience, Hoboken, N.J.

Borsuk, M. E., Stow, C. A., Reckhow, K. H., 2002. Predicting the frequency of water quality standard violations: A probabilistic approach for TMDL development. Environmental Science & Technology 36 (10), 2109–2115.

Briones, A. M., Reichardt, W., 1999. Estimating microbial population counts by 'most probable number' using Microsoft Excel. Journal of Microbiological Methods 35 (2), 157–161.

Christian, R. R., Pipes, W. O., 1983. Frequency Distributions of Coliforms in Water Distribution Systems. Applied and Environmental Microbiology 45 (2), 603–609.

Cochran, W. G., 1950. Estimation of Bacterial Densities by Means of the 'Most Probable Number'. Biometrics 6 (2), 105–116.

Crainiceanu, C. M., Stedinger, J. R., Ruppert, D., Behr, C. T., 2003. Modeling the U.S. national distribution of waterborne pathogen concentrations with application to Cryptosporidium parvum. Water Resources Research 39 (9).

de Man, J. C., 1977. MPN Tables For More Than One Test. European Journal of Applied Microbiology and Biotechnology 4 (4), 307–316.

de Man, J. C., 1983. MPN Tables, Corrected. European Journal of Applied Microbiology and Biotechnology 17 (5), 301–305.

272 Dufour, A. P., Cabelli, V. J., 1975. Membrane-Filter Procedure for Enumer-
273 ating Component Genera of Coliform Group in Seawater. Applied Microbi-
274 ology 29 (6), 826–833.
275 Dufour, A. P., Strickland, E. R., Cabelli, V. J., 1981. Membrane-Filter Method
276 for Enumerating Escherichia coli. Applied and Environmental Microbiology
277 41 (5), 1152–1158.
278 El-Shaarawi, A. H., Esterby, S. R., Dutka, B. J., 1981. Bacterial Density in
279 Water Determined by Poisson or Negative Binomial Distributions. Applied
280 and Environmental Microbiology 41 (1), 107–116.
281 Gelman, A. J., Carlin, J. B., Stern, H. S., Rubin, D. B., 2004. Bayesian Data
282 Analysis. Chapman & Hall/CRC, Boca Raton, Florida.
283 George, R., 2008. The Big Necessity. Metropolitan, New York.
284 Gleick, P. H., 1998. Water in crisis: paths to sustainable water use. Ecological
285 Applications 8 (3), 571–579.
286 Greenwood, M., Yule, G. U., 1917. On the Statistical Interpretation of Some
287 Bacteriological Methods Employed in Water Analysis. The Journal of Hy-
288 giene 16 (1), 36–54.
289 Gronewold, A. D., Borsuk, M. E., 2009. A software tool for translating deter-
290 ministic model results into probabilistic assessments of water quality stan-
291 dard compliance. Environmental Modelling & Software 24 (10), 1257–1262.
292 Gronewold, A. D., Borsuk, M. E., 2010. Improving water quality assessments
293 through a hierarchical Bayesian analysis of variability. Environmental Sci-
294 ence & Technology 44 (20), 7858–7864.
295 Gronewold, A. D., Borsuk, M. E., Wolpert, R. L., Reckhow, K. H., 2008.
296 An Assessment of Fecal Indicator Bacteria-Based Water Quality Standards.
297 Environmental Science & Technology 42 (13), 4676–4682.
298 Gronewold, A. D., Myers, L., Swall, J. L., Noble, R. T., 2010. Addressing un-
299 certainty in fecal indicator bacteria dark inactivation rates. Water Research
300 45 (2), 652–664.
301 Gronewold, A. D., Wolpert, R. L., 2008. Modeling the Relationship Between

Most Probable Number (MPN) and Colony-Forming Unit (CFU) Estimates of Fecal Coliform Concentration. Water Research 42 (13), 3327–3334.

Haas, C. N., 1989. Estimation of microbial densities from dilution count experiments. Applied and Environmental Microbiology 55 (8), 1934–1942.

Haas, C. N., Heller, B., 1988. Test of the validity of the Poisson assumption for analysis of Most-Probable-Number results. Applied and Environmental Microbiology 54 (12), 2996–3002.

Hurley, M. A., Roscoe, M. E., 1983. Automated Statistical Analysis of Microbial Enumeration by Dilution Series. Journal of Applied Bacteriology 55 (1), 159–164.

Klee, A. J., 1993. A Computer-Program For The Determination Of Most Probable Number And Its Confidence-Limits. Journal of Microbiological Methods 18 (2), 91–98.

Lunn, D. J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS-A Bayesian modelling framework: Concepts, structure, and extensibility. Statistics and Computing 10 (4), 325–337.

McBride, G. B., 2005. Using statistical methods for water quality management. Issues, problems and solutions. John Wiley & Sons Ltd Chichester, UK.

McBride, G. B., Ellis, J. C., 2001. Confidence of compliance: a Bayesian approach for percentile standards. Water Research 35 (5), 1117–1124.

McBride, G. B., McWhirter, J. L., Dalgety, M. H., 2003. Uncertainty in most probable number calculations for microbiological assays. Journal of AOAC International 86 (5), 1084–1088.

McCrady, M. H., 1915. The Numerical Interpretation of Fermentation Tube Results. Journal of Infectious Diseases 17 (1), 183–212.

Mekonnen, M. M., Hoekstra, A. Y., 2016. Four billion people facing severe water scarcity. Science Advances 2 (2), e1500323.

Messner, M. J., Wolpert, R. L., 2002. Cryptosporidium and Giardia occurrence in ICR drinking water sources - statistical analysis of ICR data. In: McGuire, M. J. (Ed.), Information Collection Rule Data Analysis. American Water

Works Association Research Foundation, Denver, CO, pp. 463–481.

Murcott, S., Keegan, M., Hanson, A., Jain, A., Knutson, J., Liu, S., Tanphanich, J., Wong, T. K., 2015. Evaluation of microbial water quality tests for humanitarian emergency and development settings. Procedia Engineering 107, 237–246.

Novotny, V., 2003. Water quality: Diffuse pollution and watershed management, 2nd Edition. John Wiley & Sons, New Jersey.

Press, S. J., 2003. Subjective and Objective Bayesian Statistics: Principles, Models, and Applications. Wiley-Interscience, Hoboken, NJ.

R core team, 2014. R: A language and environment for statistical computing. Vienna, Austria.
URL `http://www.r-project.org`

Stauber, C., Miller, C., Cantrell, B., Kroell, K., 2014. Evaluation of the compartment bag test for the detection of Escherichia coli in water. Journal of Microbiological Methods 99, 66–70.

Tillett, H. E., Coleman, R., 1985. Estimated Numbers of Bacteria in Samples from Non-Homogeneous Bodies of Water - How Should MPN and Membrane Filtration Results Be Reported? Journal of Applied Bacteriology 59 (4), 381–388.

Weiss, P., Aw, T. G., Urquhart, G. R., Galeano, M. R., Rose, J. B., 2016. Well water quality in rural Nicaragua using a low-cost bacterial test and microbial source tracking. Journal of Water and Health 14 (2), 199–207.

Woodward, R. L., 1957. How probable is the most probable number? Journal of the American Water Works Association 49 (1), 1060–1068.

Woomer, P. L., Bennett, J., Yost, R., 1990. Overcoming the Inflexibility of Most-Probable-Number Procedures. Agronomy Journal 82 (2), 349–353.

World Health Organization, 2004. Guidelines for drinking-water quality: recommendations. Tech. rep., Geneva.

Wu, J., Gronewold, A. D., Rodriguez, R. A., Stewart, J., Sobsey, M. D., 2014. Integrating quantitative PCR and Bayesian statistics in quantifying human

362 adenoviruses in small volumes of source water. Science of the Total Envi-
363 ronment 470, 255–262.

| 56mL | 30mL | 10mL | 3mL | 1mL | MPN | $q_{2.5}$ | $q_{5.0}$ | $q_{25.0}$ | $q_{75.0}$ | $q_{95.0}$ | $q_{97.5}$ | Likelihood that $c >$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | 1 | 10 | 100 | 1000 |
| 0 | 0 | 0 | 0 | 0 | 0.0 | <0.1 | <0.1 | 0.3 | 1.4 | 3.0 | 3.7 | 0.33 | 0.00 | 0.00 | 0.00 |
| 1 | 0 | 0 | 0 | 0 | 1.5 | 0.4 | 0.5 | 1.5 | 4.4 | 8.1 | 9.7 | 0.85 | 0.02 | 0.00 | 0.00 |
| 1 | 1 | 0 | 0 | 0 | 4.7 | 1.5 | 2.1 | 4.8 | 13.2 | 24.8 | 29.7 | 0.99 | 0.39 | 0.00 | 0.00 |
| 1 | 1 | 1 | 0 | 0 | 13.6 | 4.8 | 6.4 | 15.1 | 44.1 | 84.4 | 101.8 | 1.00 | 0.87 | 0.03 | 0.00 |
| 1 | 1 | 0 | 1 | 0 | 9.6 | 3.3 | 4.3 | 9.2 | 23.0 | 39.7 | 46.5 | 1.00 | 0.71 | 0.00 | 0.00 |
| 1 | 1 | 1 | 1 | 0 | 48.3 | 16.4 | 22.4 | 55.5 | 170.0 | 331.0 | 400.5 | 1.00 | 0.99 | 0.50 | 0.00 |
| 1 | 1 | 1 | 0 | 1 | 32.6 | 10.9 | 14.3 | 31.6 | 81.4 | 141.8 | 166.5 | 1.00 | 0.98 | 0.16 | 0.00 |
| 1 | 1 | 1 | 1 | 1 | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA | NA |

Table 1. Likelihood function-based interpretation of the eight highest likelihood combinations of positive (1) and negative (0) compartments for the recommended 5-compartment CBT design. Results include the MPN, quantiles of the FIB concentration normalized likelihood function, and the likelihood-based probability that $c$ exceeds numeric water quality guidelines of 1, 10, 100, and 1000 (organisms per 100 ml). Note that when all compartments are positive, the FIB concentration likelihood function is continuously increasing and therefore the MPN and FIB concentration quantiles are undefined (Gronewold et al., 2010).

| 56mL | 30mL | 10mL | 3mL | 1mL | MPN | $q_{2.5}$ | $q_{5.0}$ | $q_{25.0}$ | $q_{75.0}$ | $q_{95.0}$ | $q_{97.5}$ | Post. prob. that $c >$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | | 1 | 10 | 100 | 1000 |
| 0 | 0 | 0 | 0 | 0 | 0 | <0.1 | <0.1 | <0.1 | <0.1 | 0.5 | 0.8 | 0.02 | 0.00 | 0.00 | 0.00 |
| 1 | 0 | 0 | 0 | 0 | 1.5 | <0.1 | 0.1 | 0.4 | 2.2 | 4.9 | 6.2 | 0.51 | 0.00 | 0.00 | 0.00 |
| 1 | 1 | 0 | 0 | 0 | 4.7 | 0.6 | 0.9 | 2.5 | 8.3 | 18.1 | 22.0 | 0.94 | 0.18 | 0.00 | 0.00 |
| 1 | 1 | 1 | 0 | 0 | 13.6 | 2.8 | 3.8 | 11.1 | 49.7 | 142.5 | 189.3 | 1.00 | 0.78 | 0.10 | 0.00 |
| 1 | 1 | 0 | 1 | 0 | 9.6 | 0.6 | 0.8 | 2.5 | 8.3 | 18.9 | 23.6 | 0.94 | 0.19 | 0.00 | 0.00 |
| 1 | 1 | 1 | 1 | 0 | 48.3 | 2.8 | 3.9 | 11.0 | 50.8 | 143.6 | 191.2 | 1.00 | 0.78 | 0.10 | 0.00 |
| 1 | 1 | 1 | 0 | 1 | 32.6 | 37.4 | 70.7 | 734.7 | $8.5 \times 10^5$ | $1.2 \times 10^9$ | $1.9 \times 10^{10}$ | 1.00 | 1.00 | 0.92 | 0.70 |
| 1 | 1 | 1 | 1 | 1 | NA | 36.9 | 67.2 | 702.9 | $8.5 \times 10^5$ | $1.4 \times 10^9$ | $2.0 \times 10^{10}$ | 1.00 | 1.00 | 0.93 | 0.73 |

Table 2. Bayesian interpretation of the eight highest likelihood combinations of positive (1) and negative (0) compartments for the recommended 5-compartment CBT design. Results include the MPN, quantiles of the FIB concentration posterior probability distribution, and the posterior probability that $c$ exceeds numeric water quality guidelines of 1, 10, 100, and 1000 (organisms per 100 ml). Note that with a Bayesian interpretation, quantiles of the FIB concentration posterior probability distribution are defined when all compartments are positive, however the MPN is not defined (when all compartments are positive) because it is based on the likelihood function alone (Gronewold et al., 2010). A Bayesian interpretation of all possible combinations of positive and negative compartments is included in the Supplementary Information.

| Highest likelihood combinations of pos. (1) and neg. (0) compartments | | | | | MPN | | | $c_{95}$ | | | P($c$>100) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 56mL | 30mL | 10mL | 3mL | 1mL | * | ** | *** | * | ** | *** | * | ** | *** |
| 0 | 0 | 0 | 0 | 0 | 0.0 | 0.0 | 0.0 | 0.5 | 0.5 | 0.4 | 0.0 | 0.0 | 0.0 |
| 1 | 0 | 0 | 0 | 0 | 1.5 | 1.4 | 1.2 | 4.9 | 4.6 | 4.1 | 0.0 | 0.0 | 0.0 |
| 1 | 1 | 0 | 0 | 0 | 4.7 | 4.1 | 3.4 | 18.1 | 14.0 | 10.2 | 0.0 | 0.0 | 0.0 |
| 1 | 1 | 1 | 0 | 0 | 13.6 | 13.0 | 8.4 | 142.5 | 175.0 | 50.2 | 0.1 | 0.1 | 0.0 |
| 1 | 1 | 0 | 1 | 0 | 9.6 | 7.8 | 5.8 | 18.9 | 14.4 | 10.2 | 0.0 | 0.0 | 0.0 |
| 1 | 1 | 1 | 1 | 0 | 48.3 | 60.9 | 19.6 | 143.6 | 178.8 | 49.5 | 0.1 | 0.1 | 0.0 |
| 1 | 1 | 1 | 0 | 1 | 32.6 | 36.2 | 17.3 | $1.2 \times 10^9$ | $1.8 \times 10^9$ | $0.9 \times 10^9$ | 0.9 | 0.9 | 0.8 |
| 1 | 1 | 1 | 1 | 1 | NA | NA | NA | $1.4 \times 10^9$ | $1.8 \times 10^9$ | $1.0 \times 10^9$ | 0.9 | 0.9 | 0.8 |

Table 3. Comparison between results when there is minimal (or no) user error (*) and results with there is either moderate (**) or severe (***) user error. MPN values and $95^{th}$ percentiles of the FIB concentration ($c_{95}$) are in organisms per 100 ml. The final column indicates the posterior probability that the FIB concentration $c$ exceeds the WHO numeric water quality standard of 100 organisms per 100 ml. Note that with a Bayesian interpretation, quantiles of the FIB concentration posterior probability distribution are defined when all compartments are positive, however the MPN is not defined (when all compartments are positive) because it is based on the likelihood function alone (Gronewold et al., 2010).

**Figure 1**

**Click here to download Figure: likelihood_evolution.ps**

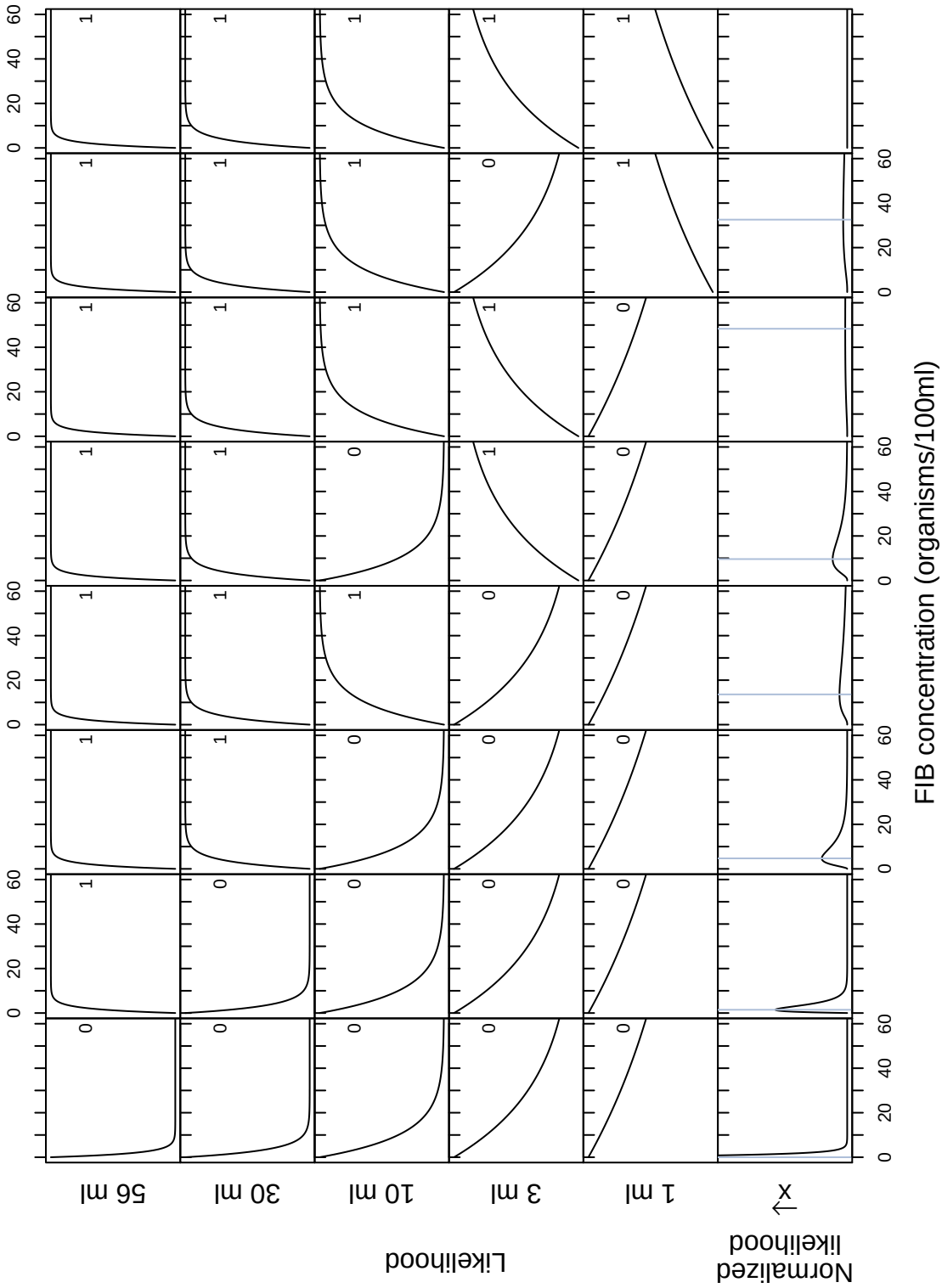


**Figure 2**

**Click here to download Figure: bayes_inf.ps**

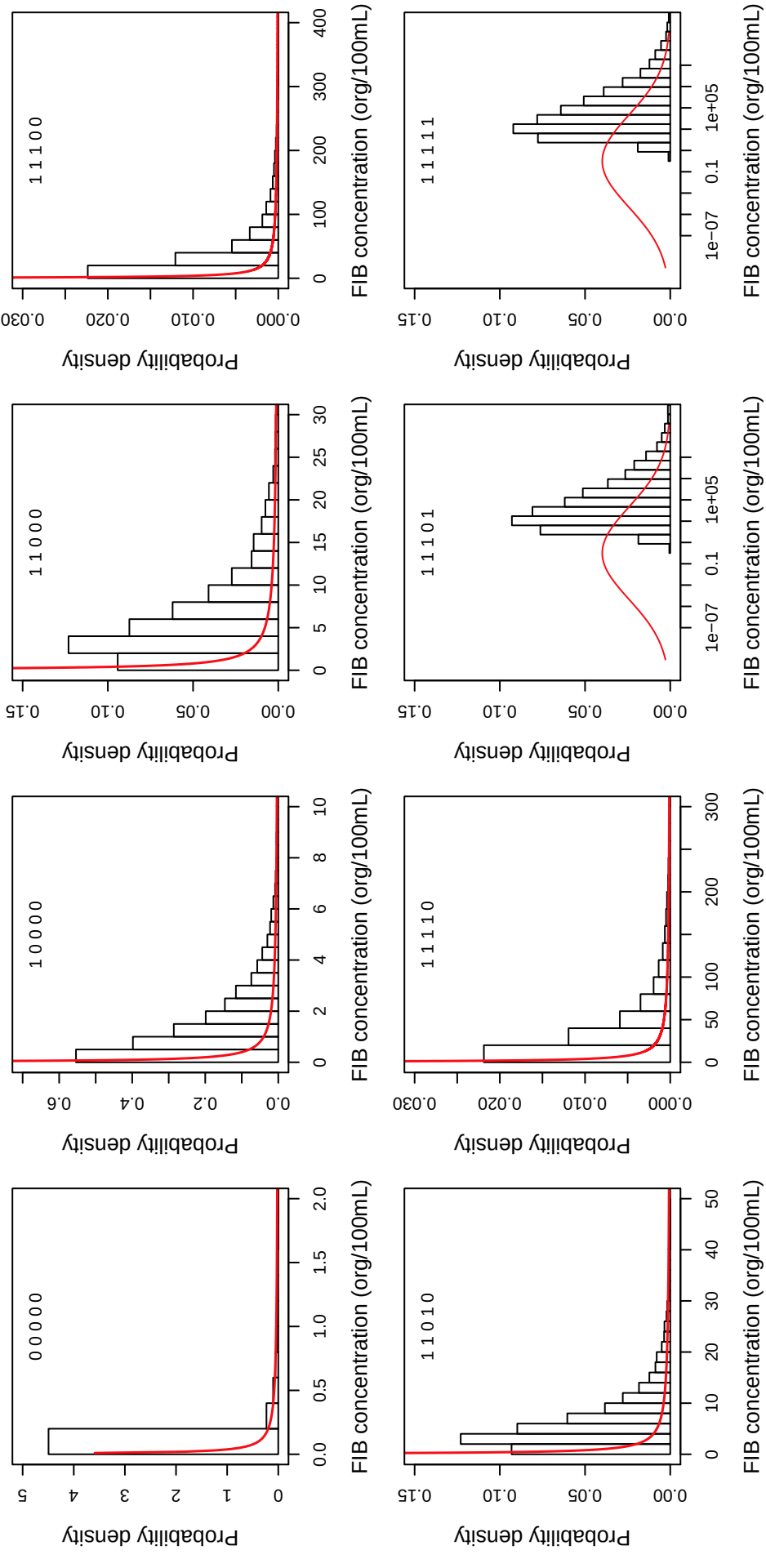
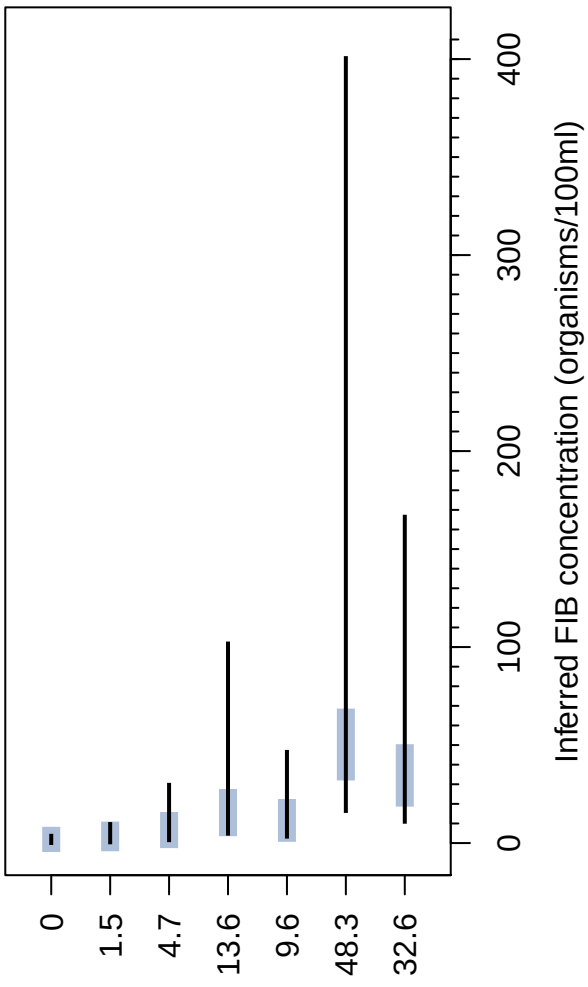


**Figure 3**

**Click here to download Figure: intervals.ps**



**Supplementary material for on-line publication only**

**Click here to download Supplementary material for on-line publication only: cbt_si_06feb2017_SUBMIT.pdf**

**LaTeX manuscript file**

**Click here to download LaTeX Source Files: cbt_design.tex**

**LaTeX bibliography file**

**Click here to download LaTeX Source Files: library.bib**