NeuroToxicology 36 (2013) 34–41

Contents lists available at SciVerse ScienceDirect

NeuroToxicology

Evaluation of microelectrode array data using Bayesian modeling as an approach to screening and prioritization for neurotoxicity testing☆

William R. LeFew [a], Emma R. McConnell [b,1], James L. Crooks [c], Timothy J. Shafer [a,*]

[a] Integrated Systems Toxicology Division, NHEERL, ORD, U.S. Environmental Protection Agency, Research Triangle Park, NC, United States
[b] Axion Biosystems, Atlanta, GA, United States
[c] Biostatistics and Bioinformatics Research Core, NHEERL, ORD, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, United States

ARTICLE INFO

Article history:
Received 6 December 2012
Accepted 16 February 2013
Available online 27 February 2013

Keywords:
Chemical screening
Neurotoxicity
Microelectrode array
Data analysis

ABSTRACT

The need to assess large numbers of chemicals for their potential toxicities has resulted in increased emphasis on medium- and high-throughput in vitro screening approaches. For such approaches to be useful, efficient and reliable data analysis and hit detection methods are also required. Assessment of chemical effects on neuronal network activity using microelectrode arrays (MEAs) has been proposed as a screening tool for neurotoxicity. The current study examined a Bayesian data analysis approach for assessing effects of a 30 chemical training set on activity of primary cortical neurons grown in multi-well MEA plates. Each well of the MEA plate contained 64 microelectrodes and the data set contains the number of electrical spikes registered by each electrode over the course of each experiment. A Bayesian data analysis approach was developed and then applied to several different parsings of the data set to produce probability determinations for hit selection and ranking. This methodology results in an approach that is approximately 74% sensitive in detecting chemicals in the training set known to alter neuronal function (23 expected positives) while being 100% specific in detecting chemicals expected to have no effect (7 expected negatives). Additionally, this manuscript demonstrates that the Bayesian approach may be combined with a previously published weighted mean firing rate approach in order to produce a more robust hit detection method. In particular, when combined with the weighted mean firing rate approach, the joint analysis produces a sensitivity of approximately 96% and a specificity of 100%. These results demonstrate the utility of a novel approach to analysis of MEA data and support the use of neuronal networks grown on MEAs as a for neurotoxicity screening approach.

Published by Elsevier Inc.

1. Introduction

The National Academies report on Toxicity Testing in the 21st Century highlighted the need to characterize the toxicity of thousands of chemicals present in the environment (NRC, 2007) to provide adequate protection of human health. As a result, there has been a substantial effort to develop rapid, cost-efficient methods to screen thousands of chemicals for their potential to cause toxicity. This effort includes new approaches to characterizing the potential for chemicals to disrupt function of the nervous system, following both acute (Novellino et al., 2011; Defranchi et al., 2011; McConnell et al., 2012), and developmental exposure (Breier et al., 2008; Radio et al., 2008; Robinette et al., 2011; Hogberg et al., 2011).

One approach that has been proposed as a screening method for neurotoxicity is the use of microelectrode array (MEA) recordings from primary cultures of neurons (for review, see Johnstone et al., 2010). Recently, several studies have demonstrated that detection of chemical effects on of neuronal network function is reproducible across different laboratories (Novellino et al., 2011) and that MEA-based assays have high specificity and selectivity (Defranchi et al., 2011; McConnell et al., 2012). These results with small sets of chemicals (20–30 compounds), indicate that testing larger numbers of chemicals using MEAs is feasible. As larger libraries of compounds are examined, particularly those where the potential actions on neuronal network activity are unknown, it will be necessary to have unbiased approaches to determine which

☆ This work was conducted as part of a Cooperative Research and Development Agreement between the U.S. EPA Office of Research and Development and Axion Biosystems (CRADA 644-11). Preparation of this document has been funded by the U.S. Environmental Protection Agency. This document has been reviewed by the National Health and Environmental Effects Research Laboratory and approved for publication. Approval does not signify that the contents reflect the views of the Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use. The R code used for this document is available by request to WRL (lefew.william@epa.gov).

* Corresponding author at: Integrated Systems Toxicology Division, MD-B105-03, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, United States. Tel.: +1 919 541 0647; fax: +1 919 541 4849.
E-mail address: Shafer.tim@epa.gov (T.J. Shafer).
[1] Current address: Nicholas School for the Environment, Duke University, Durham, NC, United States.

0161-813X/$ – see front matter. Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.neuro.2013.02.006

chemicals alter activity (“hit” detection) and prioritize them for additional testing.

To date, detection of chemical effects using MEAs has been based on changes in the mean firing rate (MFR) of the network of neurons in each array (Defranchi et al., 2011; McConnell et al., 2012). In doing so, the data that are obtained from the typically 60–64 electrodes in the array are averaged to a single value for each concentration of compound that is examined. From a pathophysiological standpoint, while changes in network firing rates may be an important indicator of neuroactivity or neurotoxicity, patterns and distributions of activity in neural networks are extremely important to physiological processes such as network formation during development, plasticity, and information sharing and distribution (Crumiller et al., 2011; Uhlhaas et al., 2009; Banerjee and Ellender, 2009). Therefore, detection of changes in distributions of activity across a network may also be important terms of screening compounds for potential neurotoxicity. One of the advantages of MEA approaches is that they allow the opportunity to record from multiple individual neurons in a network simultaneously. Thus, the practice of averaging data across all electrodes does not fully utilize the high content aspect of the data collected from MEAs and more importantly may not detect changes in other significant functional parameters.

The goal of the present study was to examine alternative approaches to analyze MEA data for hit detection and chemical prioritization. As an alternative to averaging all the data from one well into a single measure of activity (e.g. MFR), Bayesian approaches considering data from individual electrodes were examined. By utilizing Bayesian techniques, this larger, electrode-based data set was used to build firing-rate distributions by electrode for each chemical tested. Then, on a distributional basis, comparisons can be made between effects of control and chemical treatment on network activity. The resultant output may still be simplified to one or few metrics in order to simplify hit detection and prioritization, but the basis for the output is a more detailed descriptor which is derived and examined in the process. This will allow researchers to immediately extract firing rate characteristics for chemicals of interest that would have been otherwise hidden by simpler tests. To illustrate the utility of this approach with a simple example, consider the following theoretical effect of a chemical on firing rate. In the control condition, the firing rate across all electrodes in the array is represented by a Gaussian distribution (Fig. 1). Following chemical treatment, firing rates on some electrodes decrease, while it increases on other electrodes, such that the distribution across all electrodes becomes bi-modal. However, if the increases in activity offset the decreases, then it is certainly possible that the MFR of the control and treated distributions are the same. As a consequence, an automated approach that evaluates only the MFR to determine chemical “hits” would miss the differences between the distributions and hence the chemical effect, resulting in a false negative. With the suggested Bayesian approach and metric outlined below, the difference between the two distributions would be evident and easily detected.

The data set for the present analysis comes from a previous study which examined the ability of MEAs to detect changes in network function following exposure to a single concentration of 30 different chemicals (McConnell et al., 2012). The design of the experiment consisted of recording 30 min of control data followed by 30 min of data in the presence of each chemical using multi-well MEAs. Each of the 30 chemicals was assessed a minimum of three times. Hits were detected based on the ability of individual chemicals to alter the weighted mean firing rate in comparison to the vehicle control, but hits were not prioritized for further screening tests (McConnell et al., 2012).

2. Materials and methods

2.1. Data collection

A multi-well MEA system (Axion Biosystems Maestro system) was utilized to determine the ability of a training set of 30

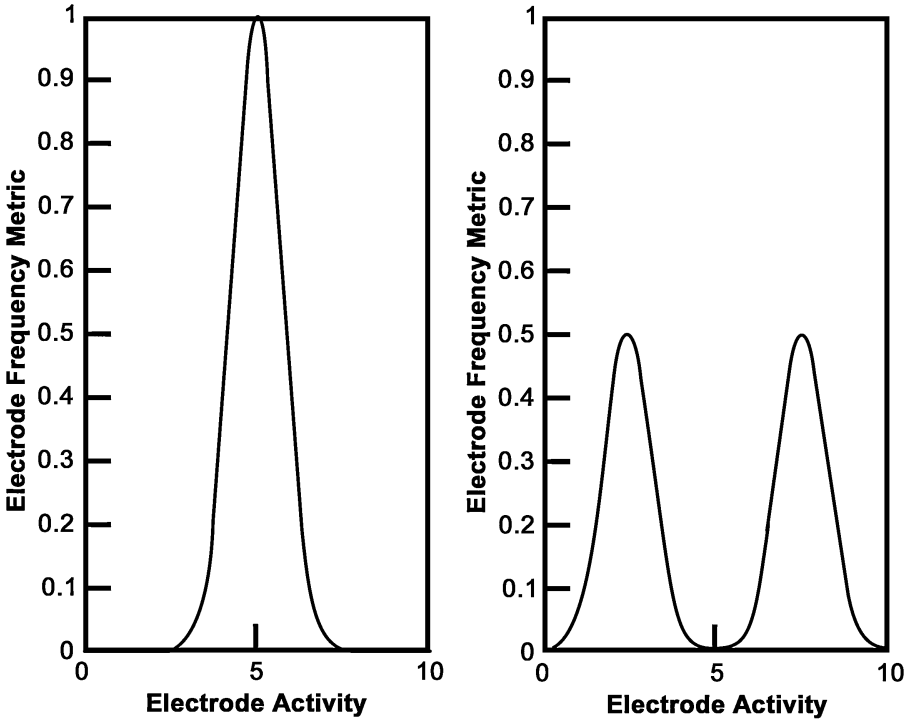

Fig. 1. These plots display two different distributions which have the same mean. The Bayesian approach will differentiate between these behaviors when simple averaging approaches will not.

Table 1
Chemical training set.[a]

Chemical positives	Chemical negatives
Bicuculline	Acetaminophen
Bifenthrin[a]	Amoxicillin
Carbaryl	Glyphosate
Chlorpyrifos oxon	Paraquat
β-Cyfluthrin[a]	Saccharin
Deltamethrin	Salicylic acid
Diazepam	D-Sorbitol
Domoic acid[b]	
Fipronil	
Fluoxetine	
Imidacloprid	
Ketamine	
Lead[b]	
L-Glutamate	
Lindane	
Methylmercury[b]	
Muscimol	
Nicotine	
Permethrin	
RDX[a]	
Trimethyltin[b]	
Valproic acid	
Verapamil	

[a] Primary cortical neuronal cultures were grown on 12 well MEA plates for a minimum of 14 days in vitro. On the day of the experiment, 33 min of baseline activity was recorded, individual wells were treated with one of the training set chemicals (50 μM, unless otherwise noted) or DMSO (vehicle control), and another 33 min of activity was recorded. Complete details on chemical purity, source and effects on mean firing rate can be found in McConnell et al. (2012). Twenty-three of the chemicals are known to disrupt nervous system function (chemical positives), while 7 of the chemicals are known for their relative lack of effects on neuronal activity at reasonable doses (chemical negatives). In subsequent tables, DMSO was used as a solvent and was not considered against correct/incorrect hit identification for each data parsing.

[b] For domoic acid, methylmercury, and trimethyltin a 10 μM concentration was used. For lead a 30 μM concentration was used.

chemicals (Table 1) to alter network activity of primary cortical cultures. Full details of culture protocols, recording conditions, chemical sources and purity are provided in McConnell et al. (2012). Briefly, the data set consists of 33 min recordings in the absence and presence of chemicals and is comprised of recordings of 64 electrodes per well from 158 total wells on 24 plates from 10 different culture preparations. Data consisted of spike counts taken on an electrode by electrode basis on one second intervals after removing the first 3 min from each condition to allow for stabilization of activity and avoid addition artifacts. Thus, 30 min of activity was analyzed in both the baseline and treated conditions.

2.2. Data analysis

The goal of the data analysis was to compare the distribution of firing rate activity on electrodes following chemical treatment to the distribution of firing rate activity on electrodes in the control condition (baseline). As behavior can vary from electrode to electrode, this Bayesian approach utilizes MEA data in greater detail by evaluating activity over all electrodes in a well rather than averaging the data from a well. The electrode firing rate of both the baseline condition and chemical treatments (includes treatment with the vehicle, DMSO) were categorized utilizing multinomial distributions and then compared to each other using Bayesian mixture analysis. This resulted in a numerical value which reflects the similarity of distribution of electrode firing rate behaviors in the treated case (the presence of chemical) to the distribution of electrode firing rate behaviors in the baseline case. This process is outlined in detail below. The computational analysis was performed in R version 2.14.0. The R code is available by request to lefew.william@epa.gov.

The spike count files generated from recordings were imported into a custom R script for analysis. The first 3 min of each data file were removed in order to let neuronal activity stabilize in the Axion Biosystems Maestro and remove addition artifacts. As with the original study (McConnell et al., 2012), data from wells that contained fewer than 10 active electrodes (active ≥5 spikes/min) were not included in analysis. Data from single electrodes were removed when observed changes in spike rate indicated noise on that channel. Only 4 electrodes were removed from 4 separate wells for this reason.

Firing rate activity was analyzed on an electrode by electrode basis rather than well by well basis. The electrodes were first sorted by chemical. For each chemical, the spike count of each electrode was used to compute a spikes/minute average for each electrode. The collection of spikes/minute averages for each electrode was then ordered from least to greatest and stored for use in analysis. For each chemical, this collection of spikes/minute averages forms a distribution.

The distribution of spikes/minute averages generated during the initial 30 min of recording characterizes the baseline distribution. This distribution of spike/minute averages was parsed in three distinct ways for the purpose of analysis. In the first case (data parsing I), the distribution was divided into two subsets, inactive electrodes and active electrodes. Inactive electrodes were defined as those electrodes that averaged fewer than 5 spikes per minute. This definition of inactivity is by convention and follows the method used in Robinette et al. (2011) and McConnell et al. (2012). This division of data is primarily concerned with the percentage of electrodes which are active/inactive and ignores the type of activity which occurs. In the second case (data parsing II), the active electrodes, as previously defined, were additionally binned into their deciles (10 quantiles) such that 11 total bins (10 quantiles of active electrodes plus the inactive electrode bin) were used in analysis. This choice of granularity was arbitrary, but tailored to the large amount of control data which was available. In order for this second method to be used in general, there must be enough bins to accurately characterize the distribution, but not so many that the bins become unreasonably narrow. Fig. 2 illustrates this separation of data into bins. The second method of analysis results in an output which accounts for both the number of active

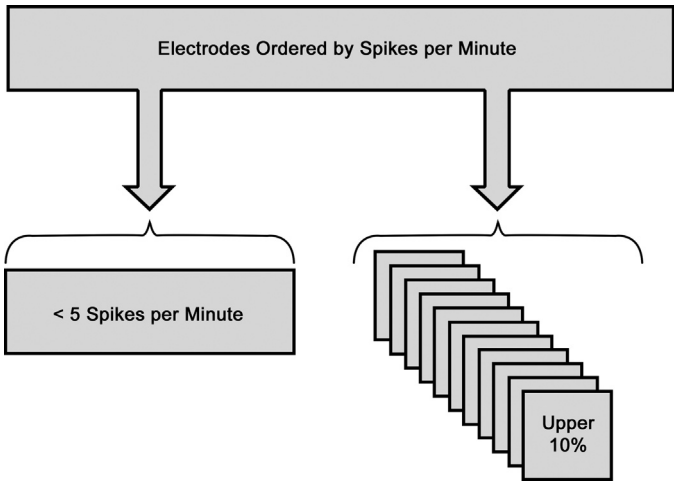

Fig. 2. This displays the way in which data were parsed. In type I parsing, the data are separated only into inactive and active electrodes. In type II parsing, the active data are additionally parsed into deciles. These deciles and the inactive electrodes comprise a type II parsing. The analysis of only the deciles comprises a type III parsing.

Table 2
Summary of the different approaches for binning electrode spikes/minute data.

	Data parsing I	Data parsing II	Data parsing III
Control[a]	Inactive and active electrodes, 2 total bins	Inactive and active electrodes, active electrodes in deciles, 11 total bins	Active electrodes in deciles, 10 total bins
Treated	Inactive and active electrodes, 2 total bins	Inactive and active electrodes, active electrodes in 10 bins bounded by the decile bounds of the control, 11 total bins	Active electrodes in 10 bins bounded by the decile bounds of the control, 10 total bins

[a] The control data is considered in two separate ways. In the first, all electrode data gathered from all control experiments are pooled and the behavior of each treated chemical's electrodes is compared to this pooled data (Type A). In the second, only electrode data gathered from the control experiment which immediately precedes the application of the treated chemical (in the same well) is utilized in comparison to the treated chemical's electrode firing rate behavior (Type B).

electrodes and the distribution of that activity. In the third case (data parsing III) only the active electrodes, binned into their deciles, were analyzed. Analysis using this third parsing of the data set results in a metric which is concerned with the distribution of spikes/minute averages and ignores the percentage of electrodes which are inactive.

The treated data were parsed in a similar way with an important exception. In the cases where deciles were used for the control chemical, those same decile bounds were used to parse the active electrodes of the treated data. This creates a distribution of active treated data electrodes within the ten decile bins that is not necessarily evenly distributed. These details are illustrated in Table 2.

The goal with any of the data parsings is to calculate the probability that the treated spikes/minute distribution is drawn from a distribution that has the same frequencies as the control, versus a distribution that does not. Roughly speaking, this probability informs as to whether the treated data is like the control data while implicitly processing the myriad of ways in which this similarity could occur. Bayesian modeling is utilized in order to generate this probability. The mathematics involved in this formulation is described below.

The alternative hypotheses, similar or not, are denoted as M_1 and M_2, respectively, and before observing the treated data it is assumed that they are equally likely, i.e. $p(M_1) = p(M_2) = 1/2$. The first alternative, M_1, would claim that the (unobserved) proportion of electrodes in each bin (also referred to as bin probabilities) underlying the treated data are the same as the (observed) proportion of electrodes in each bin of the control data, whereas the second alternative, M_2, says that all possible sets of proportions of electrodes in each bin underlying the treated data are equally likely and are not similar to the control proportions. The alternatives themselves are formally written as $p(\theta|M_1) = I(\theta = \theta_c)$ (the indicator function such that $I(\theta = \theta_c)$ returns a 1 when $\theta = \theta_c$ and a 0 otherwise) and $p(\theta|M_2) = \pi(\theta|\alpha)$ (where $\pi(\theta|\alpha)$ represents a Dirichlet distribution with shape parameter vectorα). Here, θ is the vector which contains the proportion of electrodes in each bin and θ_c is the vector of observed electrode proportions in the control data (for instance, each element in θ_c is 0.1 in the case of data parsing III as the control data has been separated into deciles). Note that the sums of the components of θ and θ_c are each 1. Furthermore, all components of α are set to one, implying that $\pi(\theta|\alpha)$ is a flat distribution over the simplex.

In order to apply a Bayesian approach utilizing the structure of the previous paragraph, the likelihood that the treated data occurs given the two alternatives above must be defined. Manipulation of Bayes' Theorem will then give a metric by which the similarity of the treated data to the control data will be evaluated. If y is the **observed** treated data, then, given the vector θ, the likelihood of y given θ follows a multinomial distribution denoted $\mathcal{L}(y|\theta)$. As θ is an unknown vector of bin probabilities, it is integrated out under the two alternatives hypotheses to yield $p(y|M_1) = \mathcal{L}(y|\theta_c)$ and $p(y|M_2) = m(y|\alpha) = \mathcal{L}(y|\theta)\pi(\theta|\alpha)/\pi(\theta|\alpha + y)$, where the last equality follows due to Bayes' Theorem and the conjugacy of Dirichlet and multinomial distributions. Finally, Bayes' Theorem is used once more to calculate λ, the posterior probability of M_1 given the treated data:

$$\lambda \cong p(M_1|y) = \frac{p(y|M_1)\,p(M_1)}{p(y|M_1)\,p(M_1) + p(y|M_2)\,p(M_2)} = \frac{\mathcal{L}(y|\theta_c)}{\mathcal{L}(y|\theta_c) + m(y|\alpha)}.$$

The practical interpretation of λ is that a value close to unity indicates that the treated distribution is similar to the control distribution, whereas a value close to zero indicates that it is not. Once λ's are generated, a threshold value delineating hits (those chemicals which are similar to the control) and misses (those chemicals which are different from the control) can be chosen. This threshold value provides a simple hit detection and prioritization approach that evaluates differences in firing rate distribution which are not necessarily detected by previously published screening techniques. As mentioned above, this method is applied in three distinct cases corresponding to the three distinct ways in which data were parsed. In data parsing I, θ_c is a two element vector which contains the proportion of control electrodes which are inactive and the proportion of control electrodes which are active. In the data parsing III, the θ_c vector is a 10 element vector with 0.1 in each entry. The treated data, y, is also a 10 element vector, but contains the numbers of electrodes which fell into each of the bins described by the bounds of the control deciles. In the cases where all data are examined, the inactive electrodes are an additional bin of their own. In data parsing II, the contribution of the active electrodes is adjusted proportionally to account for the addition of the inactive electrodes to the data set (for example, if 50% of the total electrodes were inactive, then θ_c would be a vector with first entry 0.5 and subsequent entries of 0.05).

In addition to the three data parsings discussed above, the control data is considered in two separate ways. In the first, all electrode data gathered from all control experiments are pooled and the behavior of each treated chemical's electrodes is compared to this pooled data (Type A). In the second, only electrode data gathered from the control experiment which immediately precedes the application of the treated chemical (in the same well) is utilized in comparison to the treated chemical's electrode firing rate behavior (Type B). This methodology was applied to all thirty chemicals.

3. Results

As described in Section 2, computations were run on three different parsings of data and two separate embodiments of the control data. The results, in the form of λ values, are displayed and hits are counted based on a $\lambda < 0.5$ threshold. Although the selection of this level (0.5) is arbitrary, λ values less than 0.5

Table 3
Results for type IA data parsing. This approach correctly identified 4/7 negatives and 18/23 positives.

Chemicals	Inactive electrodes	Total electrodes	λ_{IA}	Correct/incorrect classification
Lindane	146	256	0.927921	Incorrect
D-Sorbitol	111	192	0.917572	Correct
Glyphosate	190	320	0.917008	Correct
L-Glutamate	149	252	0.916125	Incorrect
Chlorpyrifos oxon	156	256	0.868251	Incorrect
Valproic acid	110	180	0.866228	Incorrect
Saccharin	137	256	0.857913	Correct
DMSO	629	1149	0.851351	=
Amoxicillin	133	256	0.741524	Correct
RDX	162	254	0.595614	Incorrect
Paraquat	165	256	0.471512	Incorrect
Salicylic acid	127	256	0.370538	Incorrect
Acetaminophen	90	192	0.139478	Incorrect
Nicotine	123	256	0.130431	Correct
Methylmercury	252	384	0.061585	Correct
Fipronil	282	428	0.024053	Correct
Carbaryl	179	256	0.002508	Correct
Permethrin	226	318	4.29E−05	Correct
Cyfluthrin	231	320	4.20E−06	Correct
Deltamethrin	232	320	2.17E−06	Correct
Trimethyltin	189	254	1.55E−06	Correct
Bifenthrin	275	383	6.05E−07	Correct
Bicuculline	724	1088	1.07E−07	Correct
Imidacloprid	95	256	8.91E−09	Correct
Lead acetate	202	256	4.89E−11	Correct
Verapamil	208	256	7.09E−14	Correct
Diazepam	240	256	3.15E−37	Correct
Fluoxetine	241	256	2.81E−38	Correct
Ketamine	208	214	8.77E−40	Correct
Muscimol	189	189	4.03E−44	Correct
Domoic acid	192	192	7.71E−45	Correct

In this approach, all control data were combined and data for each chemical was compared to the combined control distribution.

Table 4
Results for type IB data parsing. This approach correctly identified 7/7 negatives, 13/23 positives.

Chemicals	Inactive electrodes	Total electrodes	λ_{IB}	Correct/incorrect classification
DMSO	629	1147	0.963988	=
Bifenthrin	275	379	0.933236	Incorrect
Glyphosate	190	320	0.932172	Correct
Paraquat	165	256	0.929869	Correct
Imidacloprid	95	254	0.929162	Incorrect
RDX	162	254	0.927663	Incorrect
Amoxicillin	133	256	0.927606	Correct
Nicotine	123	256	0.927079	Incorrect
Saccharin	137	256	0.922824	Correct
Salicylic acid	127	256	0.922687	Correct
Acetaminophen	90	191	0.915551	Correct
Lead acetate	202	255	0.91521	Incorrect
Valproic acid	110	179	0.910513	Incorrect
Lindane	146	234	0.862176	Incorrect
Permethrin	226	310	0.849595	Incorrect
Chlorpyrifos oxon	156	256	0.832451	Incorrect
D-Sorbitol	111	192	0.789307	Correct
Bicuculline	724	1087	0.685641	Incorrect
Carbaryl	179	256	0.460267	Correct
Methylmercury	252	382	0.329281	Correct
Cyfluthrin	231	319	0.173167	Correct
L-Glutamate	149	230	0.000801	Correct
Deltamethrin	232	317	9.31E−06	Correct
Trimethyltin	189	254	4.04E−06	Correct
Verapamil	208	256	1.39E−24	Correct
Fipronil	282	426	2.17E−25	Correct
Diazepam	240	256	2.11E−27	Correct
Domoic acid	192	192	1.46E−33	Correct
Ketamine	208	213	2.93E−37	Correct
Fluoxetine	241	256	4.08E−38	Correct
Muscimol	189	189	2.40E−59	Correct

In this approach, each chemical's control data were combined and data for each chemical was compared to the combined control distribution.

indicate that there is greater than chance probability that the baseline and treated distributions differ. Hits are compared to the expected results found in Table 1, wherein 23 of the chemicals are known to be neuroactive/neurotoxic and were expected to alter network firing activity, while 7 chemicals are not generally neurotoxic and not expected to alter network firing activity. DMSO was used as a solvent and was not considered against correct/incorrect hit identification for each data parsing. The use of literature compounds with known effects on neuronal activity helps to establish proof-of-concept that a test method can reliably detect chemical induced changes (Crofton et al., 2010). Overall accuracy is discussed in a manner similar to McConnell et al. (2012); this paper used a 14% change in weighted mean firing rate (wMFR) as a hit threshold and correctly classified 7/7 negative and 20/23 positive compounds from Table 1. Nicotine, imidacloprid, and bifenthrin were the three chemicals which were not correctly classified using this method.

Table 3 displays the results from processing data parsing IA. When compared to the expected results, this data parsing yields correct classification for approximately 57% (4/7) of the chemical negatives and approximately 78% (18/23) of the chemical positives. Nicotine, imidacloprid, and bifenthrin were all classified correctly as positives. By contrast, lindane, chlorpyrifos oxon and RDX were mis-classified as negatives and paraquat, salicylic acid and acetaminophen were mis-classified as positives.

Table 4 displays the results from processing data parsing IB. When compared to the training results, this data parsing yields correct classification for 100% (7/7) of the chemical negatives and approximately 56% (13/23) of the chemical positives. Nicotine, imidacloprid, and bifenthrin, as well as several other neuroactive compounds, were classified incorrectly as negatives.

Table 5 displays the results from processing data parsing IIA. When compared to the expected responses for the training set chemicals, data parsing IIA yields correct classification for 100% (7/7) of the chemical negatives and approximately 74% (17/23) of the chemical positives. Imidacloprid and bifenthrin were classified correctly as positives. Nicotine was classified incorrectly as a negative.

Table 6 displays the results from processing data parsing IIB. When compared to the training results, data parsing IIB yields correct classification for 100% (7/7) of the chemical negatives and approximately 39% (9/23) of the chemical positives. Nicotine, imidacloprid, and bifenthrin were classified incorrectly as negatives.

Table 7 displays the results from processing data parsing IIIA yields correct classification for 100% (7/7) of the chemical negatives and approximately 61% (14/23) of the chemical positives. Imidacloprid was classified correctly as a positive. Nicotine and bifenthrin were classified incorrectly as negatives.

Table 8 displays the results from processing data parsing IIIB yields correct classification for 100% (7/7) of the chemical negatives and approximately 52% (12/23) of the chemical positives. Nicotine, imidacloprid, and bifenthrin were classified incorrectly as negatives.

The results of the previous paragraphs show that parsings of Type A are superior to those of Type B. Recall that Type A parsings utilize all available electrode control data in order to make an assessment of a chemical treatment as opposed to Type B which utilizes only that electrode control data which was both in the same well and immediately preceded the chemical treatment. Additionally, parsings of Type A were more likely to categorize nicotine, imidacloprid, and bifenthrin correctly.

Table 5
Results for data parsing IIA. This approach correctly identified 7/7 negatives and 17/23 positives.

Chemicals	Total electrodes	Active electrodes	λ_{IIA}	Correct/incorrect classification
DMSO	1149	520	0.999999995	=
Valproic acid	180	70	0.999999899	Incorrect
Glyphosate	320	130	0.999999597	Correct
Paraquat	256	91	0.99999882	Correct
Amoxicillin	256	123	0.999998751	Correct
D-Sorbitol	192	81	0.999997929	Correct
Methylmercury	384	132	0.99999551	Incorrect
Saccharin	256	119	0.99999141	Correct
Salicylic acid	256	129	0.999962168	Correct
Acetaminophen	192	102	0.99979633	Correct
Nicotine	256	133	0.999323564	Incorrect
Chlorpyrifos oxon	256	100	0.998532056	Incorrect
Deltamethrin	320	88	0.993323898	Incorrect
Trimethyltin	254	65	0.79420173	Incorrect
Permethrin	318	92	0.343169127	Correct
Bifenthrin	383	108	0.254943284	Correct
Cyfluthrin	320	89	0.121271038	Correct
RDX	254	92	4.72E−05	Correct
Bicuculline	1088	364	3.07E−06	Correct
Lead acetate	256	54	2.19E−06	Correct
Carbaryl	256	77	2.21E−07	Correct
Lindane	256	110	4.58E−08	Correct
Imidacloprid	256	161	2.44E−08	Correct
L-Glutamate	252	103	4.13E−13	Correct
Verapamil	256	48	3.71E−22	Correct
Fipronil	428	146	4.51E−26	Correct
Diazepam	256	16	1.84E−27	Correct
Muscimol	189	0	4.52E−30	Correct
Ketamine	214	6	1.77E−30	Correct
Domoic acid	192	0	9.93E−31	Correct
Fluoxetine	256	15	6.93E−33	Correct

In this approach, all control data were combined and data for each chemical was compared to the combined control distribution.

Table 6
Results for data parsing IIB. This approach correctly identified 7/7 negatives, 9/23 positives.

Chemicals	Total electrodes	Active electrodes	λ_{IIB}	Correct/incorrect classification
DMSO	1149	518	0.999999993	=
Bifenthrin	383	104	0.999999986	Incorrect
Glyphosate	320	130	0.999999773	Correct
Salicylic acid	256	129	0.999999221	Correct
Lead acetate	256	53	0.999998286	Incorrect
Amoxicillin	256	123	0.999998215	Correct
Imidacloprid	256	159	0.999997834	Incorrect
D-Sorbitol	192	81	0.999997367	Correct
Paraquat	256	91	0.999996657	Correct
Valproic acid	180	69	0.999996236	Incorrect
Nicotine	256	133	0.999993557	Incorrect
Saccharin	256	119	0.999992921	Correct
Permethrin	318	84	0.999991751	Incorrect
Acetaminophen	192	101	0.999982792	Correct
Bicuculline	1088	363	0.999978445	Incorrect
RDX	254	92	0.999919545	Incorrect
Deltamethrin	320	85	0.999737773	Incorrect
Carbaryl	256	77	0.999724969	Incorrect
Lindane	256	88	0.999673708	Incorrect
Cyfluthrin	320	88	0.9996541	Incorrect
Trimethyltin	254	65	0.988711772	Incorrect
Methylmercury	384	130	0.894053887	Incorrect
Chlorpyrifos oxon	256	100	0.00729113	Correct
L-Glutamate	252	81	0.00597238	Correct
Diazepam	256	16	6.06E−08	Correct
Domoic acid	192	0	4.28E−10	Correct
Ketamine	214	5	4.85E−15	Correct
Fluoxetine	256	15	7.62E−18	Correct
Verapamil	256	48	1.29E−18	Correct
Muscimol	189	0	7.52E−25	Correct
Fipronil	428	144	2.85E−28	Correct

In this approach, each chemical's control data were combined and data for each chemical was compared to the combined control distribution.

4. Discussion

The approach described in this manuscript addresses the challenge of MEA data analysis through the design of a methodology that generates concise metrics for initial rapid screening ability after which more detailed analysis is accessible to the researcher via the detailed comparison of distributions. As with previous work on this subject (McConnell et al., 2012), the initial metric can be thresholded in order to screen rapidly chemicals.

Bayesian techniques were applied to several parsings of the available data set. As discussed in Section 3, the formulations which utilized all available control data performed best. Additionally, several formulations (IA, IIA, and IIIA) were able to correctly classify some or all of three chemicals that previously were classified as false negatives. However, as discussed below, this typically came at a cost, as all approaches had an overall lower rate of correct classifications of positives than use of wMFR.

The best overall performing analysis, formulation IIA, yielded a perfect classification rate for negatives, an approximately 74% classification rate for positives, and was able to correctly classify imidacloprid and bifenthrin. However, chlorpyrifos oxon, deltamethrin, trimethyltin and methylmercury, all well known neurotoxicants, were not detected by this method. Valproic acid was also not detected; however, the concentration of valproic acid used in the McConnell et al. (2012) experiments (50 μM) was much lower than had previously been reported (300–700 μM, Gross et al., 1995) to alter activity in MEAs. Thus, it may well be that under these experimental conditions, effects of valproic acid are at the threshold of detection, and depending on the method utilized, may or may not be considered a hit.

The weighted mean firing rate method discussed by McConnell et al. (2012) had a perfect classification rate for negatives and an approximately 87% (20/23) correct classification rate for the positives (missing nicotine, imidacloprid, and bifenthrin). Five out of the six approaches considered here were able to classify correctly 100% of the negative compounds. Thus, these results indicate that multi-well MEAs have a high specificity (correct identification of negatives), but that sensitivity (correct identification of positives) although lower, is also excellent. This is consistent with both McConnell et al. (2012) (100% specificity; 87% sensitivity) and Defranchi et al. (2011) (86% specificity; 77% sensitivity). As both methods (wMFR; Bayesian) produce simple metrics, it is interesting to point out that a combined metric (OR operator; Table 9) produces 100% specificity (a perfect classification rate for negatives) and an approximately 96% sensitivity (classification rate for positives (nicotine being the only false negative)).

Empirically, it is reasonable that combination of the two approaches using an OR operator will detect a greater number of positive chemicals, because each approach (wMFR versus Bayesian) assesses the ability of chemicals to alter spontaneous activity in different ways. Overall levels of activity (without consideration of activity distribution) are assessed using the wMFR approach, while the distribution of activity (as opposed to the overall level) is assessed with the Bayesian approaches. It is possible that chemicals may affect one of these activity measures to a greater degree than the other. For toxicity screening, a false negative response is undesirable, as the goal is to identify as many toxic compounds as possible. Because of this, the combination of the Bayesian approach with wMFR approach is potentially very powerful. Testing of larger chemical libraries that include both

Table 7
Results for IIIA. This approach correctly identified 7/7 negatives, 14/23 positives.

Chemicals	Total electrodes	Active electrodes	λ_{IIIA}	Correct/ incorrect classification
DMSO	1149	520	0.999997	=
Deltamethrin	320	88	0.999883	Incorrect
Amoxicillin	256	123	0.999774	Correct
Valproic acid	180	70	0.999767	Incorrect
Methylmercury	384	132	0.999638	Incorrect
Acetaminophen	192	102	0.999159	Correct
Paraquat	256	91	0.999114	Correct
Salicylic acid	256	129	0.99907	Correct
Glyphosate	320	130	0.998756	Correct
Nicotine	256	133	0.996753	Incorrect
D-Sorbitol	192	81	0.995814	Correct
Saccharin	256	119	0.995659	Correct
Trimethyltin	254	65	0.994603	Incorrect
Bifenthrin	383	108	0.988109	Incorrect
Diazepam	256	16	0.881458	Incorrect
Cyfluthrin	320	89	0.818532	Incorrect
Permethrin	318	92	0.70102	Incorrect
Lead acetate	256	54	0.39858	Correct
Imidacloprid	256	161	0.310501	Correct
Chlorpyrifos oxon	256	100	0.215461	Correct
Ketamine	214	6	0.029505	Correct
Bicuculline	1088	364	0.01595	Correct
Fluoxetine	256	15	0.000201	Correct
RDX	254	92	4.47E−08	Correct
Carbaryl	256	77	2.52E−08	Correct
Lindane	256	110	2.15E−11	Correct
Verapamil	256	48	2.94E−14	Correct
L-Glutamate	252	103	1.49E−16	Correct
Fipronil	428	146	1.36E−27	Correct
Muscimol	189	0	0	Correct
Domoic acid	192	0	0	Correct

In this approach, all control data were combined and data for each chemical was compared to the combined control distribution.

Table 8
Results for data parsing IIIB. This approach correctly identified 7/7 negatives and 12/23 positives.

Chemicals	Total electrodes	Active electrodes	λ_{IIIB}	Correct/incorrect classification
DMSO	1149	518	0.999984	=
Imidacloprid	256	159	0.999879	Incorrect
Salicylic acid	256	129	0.999705	Correct
Glyphosate	320	130	0.999469	Correct
Bifenthrin	383	104	0.999075	Incorrect
Amoxicillin	256	123	0.99905	Correct
Nicotine	256	133	0.998186	Incorrect
Deltamethrin	320	85	0.997127	Incorrect
Acetaminophen	192	101	0.996908	Correct
D-Sorbitol	192	81	0.996553	Correct
Saccharin	256	119	0.994643	Correct
Valproic acid	180	69	0.992112	Incorrect
Paraquat	256	91	0.97618	Correct
Trimethyltin	254	65	0.923564	Incorrect
Diazepam	256	16	0.916809	Incorrect
Bicuculline	1088	363	0.880884	Incorrect
Permethrin	318	84	0.801544	Incorrect
RDX	254	92	0.700682	Incorrect
Lead acetate	256	53	0.625773	Incorrect
Lindane	256	88	0.609776	Incorrect
Carbaryl	256	77	0.546984	Incorrect
Cyfluthrin	320	88	0.487727	Correct
Ketamine	214	5	0.050545	Correct
Methylmercury	384	130	0.007305	Correct
Fluoxetine	256	15	0.001627	Correct
L-Glutamate	252	81	0.000489	Correct
Chlorpyrifos oxon	256	100	5.91E−06	Correct
Verapamil	256	48	3.35E−10	Correct
Fipronil	428	144	1.35E−18	Correct
Domoic acid	192	0	0	Correct
Muscimol	189	0	0	Correct

In this approach, each chemical's control data were combined and data for each chemical was compared to the combined control distribution.

Table 9
Results for data parsing IIA, the wMFR classification method and the joint classification method using both IIA and wMFR with the OR operator. This approach correctly identified 7/7 negatives and 22/23 positives.

Chemicals	IIA classification	wMFR classification	Blended OR classification
DMSO	=	=	=
Valproic acid	Incorrect	Correct	Correct
Glyphosate	Correct	Correct	Correct
Paraquat	Correct	Correct	Correct
Amoxicillin	Correct	Correct	Correct
D-Sorbitol	Correct	Correct	Correct
Methylmercury	Incorrect	Correct	Correct
Saccharin	Correct	Correct	Correct
Salicylic acid	Correct	Correct	Correct
Acetaminophen	Correct	Correct	Correct
Nicotine	Incorrect	Incorrect	Incorrect
Chlorpyrifos oxon	Incorrect	Correct	Correct
Deltamethrin	Incorrect	Correct	Correct
Trimethyltin	Incorrect	Correct	Correct
Permethrin	Correct	Correct	Correct
Bifenthrin	Correct	Incorrect	Correct
Cyfluthrin	Correct	Correct	Correct
RDX	Correct	Correct	Correct
Bicuculline	Correct	Correct	Correct
Lead acetate	Correct	Correct	Correct
Carbaryl	Correct	Correct	Correct
Lindane	Correct	Correct	Correct
Imidacloprid	Correct	Incorrect	Correct
L-Glutamate	Correct	Correct	Correct
Verapamil	Correct	Correct	Correct
Fipronil	Correct	Correct	Correct
Diazepam	Correct	Correct	Correct
Muscimol	Correct	Correct	Correct
Ketamine	Correct	Correct	Correct
Domoic acid	Correct	Correct	Correct
Fluoxetine	Correct	Correct	Correct

positive and negative compounds is needed to confirm this hypothesis.

There are also other uses of the information from the Bayesian approach that could be of value in analysis of data from screening assays. For example, criteria for data acceptance/rejection could be based, in whole or part, on the similarity of activity in each individual well from a "run" of an assay to the population activity from historical controls. This would allow for an unbiased approach for determining when data from each well of a multi-well MEA would be utilized for analysis. Another use of the approach would be prioritization for subsequent testing of compounds in other in vitro or in vivo assays. Compounds with λ values further from the hit threshold (0.5, in this case) would receive higher priority than those that reached the hit threshold, but were closer to the threshold value. For example, in data parsing IIA, further testing of lindane (λ_{ALL} = 4.58E−08) would be a higher priority than for bifenthrin (λ_{ALL} = 0.25).

In conclusion, Bayesian analysis of MEA data is a novel approach to evaluation of MEA data for neurotoxicity screening. This approach can account for altered network activity following chemical treatment generated by evaluating the distribution of spike activity and comparing it to the distribution present in the baseline. It should be noted that this methodology can be easily extended to make comparisons between effects of different chemicals, in addition to the comparison to baseline made here. Such a comparison may serve as another sorting mechanism to examine the variety of ways in which the effects of chemicals may disrupt neuronal network function. While there are many variations to this approach, in all cases training data can be used to determine the best suited embodiment for screening purposes. Additionally, Bayesian analysis of MEA data may be combined with other hit detection methods in order to produce more robust

results. In particular, methods that would detect the type of effects which chemicals like nicotine display would be a useful supplement to the Bayesian (and wMFR) methodology. Such techniques would supplement the tools discussed in this manuscript and provide a more complete screening methodology.

Conflict of interest statement

None declared.

Acknowledgements

The authors are grateful for the comments and useful suggestions provided by Drs William Mundy and Lyle Burgoon on an earlier draft of this manuscript.

References

Banerjee A, Ellender TJ. Oscillations in the developing cortex: a mechanism for establishing and synchronizing an early network? J Neurosci 2009;29: 15029–30.

Breier J, Radio N, Mundy WR, Shafer TJ. Development of a high-throughput assay for assessing antiproliferative compounds using human neuroprogenitor cells. Toxicol Sci 2008;105:119–33.

Crofton K, Mundy W, Lein P, Bal-Price A, Coecke S, Seiler A, et al. Developmental neurotoxicity testing: recommendations for developing alternative methods for the screening and prioritization of chemicals. Altern Animal Exp 2010;28(1): 9–15.

Crumiller M, Knight B, Yu Y, Kaplan E. Estimating the amount of information conveyed by a population of neurons. Front Neurosci 2011;5:90.

Defranchi E, Novellino A, Whelan M, Vogel S, Ramirez T, van Ravenzwaay B, et al. Feasibility assessment of micro-electrode chip assay as a method of detecting neurotoxicity *in vitro*. Front Neuroeng 2011;4(6):1–12.

Gross GW, Rhoades BK, Azzazy HME, Wu M-C. The use of neuronal networks on multielectrode arrays as biosensors. Biosens Bioelectron 1995;10:553–67.

Hogberg HT, Sobanski T, Novellino A, Whelan M, Weiss DG, Bal-Price AK. Application of micro-electrode arrays (MEAs) as an emerging technology for developmental neurotoxicity: evaluation of domoic acid-induced effects in primary cultures of rat cortical neurons. Neurotoxicology 2011;32:158–68.

Johnstone AF, Gross GW, Weiss DG, Schroeder OH, Gramowski A, Shafer TJ. Microelectrode arrays: a physiologically based neurotoxicity testing platform for the 21st century. Neurotoxicology 2010;31(4):331–50.

McConnell ER, McClain MA, Ross J, LeFew WR, Shafer TJ. Evaluation of multi-well microelectrode arrays for neurotoxicity screening using a chemical training set. Neurotoxicology 2012;33(5):1048–57.

Novellino A, Scelfo B, Palosaari T, Price A, Sobanski T, Shafer T, et al. Development of micro-electrode array based tests for neurotoxicity: assessment of interlaboratory reproducibility with neuroactive chemicals. Front Neuroeng 2011;4(4):1–14.

NRC. Toxicity testing in the twenty-first century: a vision and a strategy. Washington, DC: The National Academies Press; 2007.

Radio N, Breier J, Shafer TJ, Mundy WR. Development of a high-throughput assay for assessing neurite outgrowth. Toxicol Sci 2008;105:106–18.

Robinette BL, Harrill JA, Mundy WR, Shafer TJ. *In vitro* assessment of developmental neurotoxicity: use of microelectrode arrays to measure functional changes in neuronal network ontogeny. Front Neuroeng 2011;4:1.

Uhlhaas PJ, Pipa G, Lima B, Melloni L, Neuenschwander S, Nikolić. et al. Neural synchrony in cortical networks: history, concepts and current status. Front Integr Neurosci 2009;3:17.