Data in Brief 54 (2024) 110380

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Dataset of in vitro measured chemicals neurotoxicity

Seweryn Ulaszek[a,b,*], Bartek Lisowski[a], Sebastian Polak[a,c]

[a] Chair of Pharmaceutical Technology and Biopharmaceutics, Faculty of Pharmacy, Jagiellonian University Medical College, Medyczna 9, 30-688 Kraków, Poland
[b] Doctoral School of Medical and Health Sciences, Jagiellonian University Medical College, Kraków, Poland
[c] Certara UK Ltd. (Simcyp Division), 1 Concourse Way, Sheffield S1 2BJ, UK

ARTICLE INFO

Article history:
Received 11 January 2024
Revised 21 March 2024
Accepted 27 March 2024
Available online 2 April 2024

Dataset link: Dataset of in-vitro measured chemicals neurotoxicity. (Original data)
Dataset link: Dataset of in-vitro measured chemicals neurotoxicity. (Original data)

Keywords:
Computational toxicology
Toxicology
Neurotoxicity
3rs
Animal-free
In silico

ABSTRACT

To understand and describe neurotoxicity mechanistically, we must first understand the processes and responses that occur within neuronal cell systems after the administration of a chemical. The dataset we present is a collection of experimental results from the literature that comprises various neurotoxic endpoints in human-derived in vitro models, allowing for easy data analysis. Currently available and free databases such as the EPA's ToxCast, which focuses on forecasting toxic health risks, are created by collecting reports on cytotoxicity testing and creating mathematical fits that could help predict the effects of a given chemical on various types of cells. We, in contrast, provide a smaller, raw, and heterogeneous dataset created solely of results on human-derived cell models that not only summarises the cytotoxic effects of certain substances but also creates a possibility for analysing the significance of the experimental set-up for the prediction of outcome.

* Corresponding author.
E-mail address: seweryn.ulaszek@uj.edu.pl (S. Ulaszek).

https://doi.org/10.1016/j.dib.2024.110380

Specifications Table

Subject	Toxicology
Specific subject area	*The neurotoxicity measured* in vitro *using human-derived cells under specified conditions and with different endpoints, such as metabolic activity, electrical activity, and others.*
Data format	Filtered, Digitised
Type of data	Table
Data collection	Data were collected manually by digitising plots published in scientific articles on in vitro neurotoxicity. In rare cases, data were copied if available as a text or table. The entire process was carried out exclusively through computer-based operations and specialised software – "Web Plot Digitizer version 4.6" [1]. The dataset consists of values obtained from signal detection by fluorescence and absorbance readers, flow cytometers, and microelectrode arrays (MEA).
Data source location	*Chair of Pharmaceutical Technology and Biopharmaceutics, Faculty of Pharmacy. Jagiellonian University Medical College, Medyczna 9, 30–688 Kraków, Poland*
Data accessibility	Repository name: Mendeley Data identification number: doi: 10.17632/9sz8c8dhjj.1 Direct URL to data: https://data.mendeley.com/datasets/9sz8c8dhjj/1. Data is also available on 'Tox-portal.com', together with data on the cardiotoxicity of drugs and drug-drug interactions [2,3].

1. Value of the Data

- The collected dataset comprises records for experiments with various endpoints, each explaining different effects of a chemical compound on the cells; all endpoints are assumed to have an effect on either cell death or loss of functions (i.e., firing rate) specific to neuronal cells.
- Each record contains information about the experimental conditions, so that it can be used for comparison between the results of different experiments as well as for inferring about the mechanism of toxicity.
- Data provide modellers with an easily accessible quantitative resource and an overview of potential modelling endpoints.
- The quantitative nature of the collected data offers new possibilities to determine toxicity mechanisms and allows for accurate comparison of results obtained in similar assays and under specified conditions by different research groups.
- The dataset reports results obtained using human-derived cells; the rationale behind this approach was the assumption that human-derived cell models are the best possible in vitro systems, allowing for extrapolation to humans.

2. Background

The main reason for gathering this dataset is to explore the possibilities of understanding neurotoxicity at a cellular level and at the level of a cell culture, and to help develop in silico computational toxicology models by summarising and categorising the results of the last decades of in vitro experiments present in the available literature. The results on this matter are presented by various authors; however, these results were obtained experimentally using certain individually defined methods, cell models, or endpoints. These experiments rarely focus solely on the chemicals' neurotoxicity but rather on a single endpoint. In the current era of information gathering and possibilities arising from data modelling, a comprehensive dataset merging and categorising all the different approaches published by various groups is potentially of great value. Unlike large and already well developed toxicological databases such as the EPA's ToxCast merged with predefined generic modelling tools like 'tcpl' [https://CRAN.R-project.org/package=tcpl], our main goal is not to summarise cytotoxicity but rather to provide data to predict the outcome of new neurotoxicity experiments based on results obtained previously in similar experimental setups (i.e., with different compounds or cell models) [4,5]. This approach might help to understand neurotoxicity and guide better experiment design in the future. Except for that,

we consider creating such a dataset valuable as it consists of already published heterogeneous data that, if not collected and merged, might be omitted in new projects willing to use it. Obtaining raw files from the authors of the articles in which the results were originally presented might be difficult due to various reasons, such as loss of raw data files, change of correspondence address, or simply lack of response from the person responsible for the storage of files [6,7]. Since the most common way to present the toxicity assessment results is by plot rather than numbers, we had to implement a digitisation tool for collecting the data. According to the applied methodology, we first requested the original raw data presented in the analysed publications from the authors without a single positive response regarding sharing the data. Due to the nature of the experiments, most of the results are presented as the control probe baseline value change in response to the application of the chemical of interest.

3. Data Description

The dataset is stored as tables in three separated commas (that is,.csv extension) files representing different endpoints and methods utilised to do the measurement, namely: 'Metabolic activity', 'Microelectrode array (MEA)" and "Others". Results are presented as means, unless otherwise stated.

The criteria for the inclusion of the data were:

- utilisation of only human-derived neuronal cultures as a cellular model,
- substance concentration and/or substance incubation time response,
- use of a pure single substance in a liquid solution; no extracts, third-compound-contaminated solutions, or heterogeneous solutions were included,
- the presence of detailed information on experimental set up in the article of interest such as: incubation time, cell density, differentiation status, type of assay, and absorbance wavelength. If just one piece of this information was missing or was not precisely described, the article was still included.

Dose response and time of incubation with a chosen drug were chosen as crucial parameters as they are expected to affect the cytotoxic effect the most. The choice of experimental methods presented in the collated datasets is a result of the adopted holistic approach. Rather than trying to track every reported modification of the in vitro cell model, we decided to gather results on the final toxic effect specific to the chemical of interest. Colorimetric assays gathered in the "Metabolic activity" and "Others" categories are those that track the activity (or lack of it) of certain enzymes, proteins, or metabolites characteristic of apoptotic, necrotic, and other death or death-approaching cell states. Because these assays are not specific for neurological activity and can be applied to other types of cell models representing other tissues and organs, their main advantages include widespread and common utilisation in cytotoxicity testing, which is why they constitute the main share of the set. On the other hand, a less popular test, i.e., the microelectrode array assay (MEA), was chosen for its ability to measure changes in the electrical activity of neurones. The lower number of records for this assay comes from the lesser popularity of MEA, probably caused by the high price of equipment and application, which is narrowed predominantly to neuronal and cardiac cell cultures. Although there have been a small number of reports with MEA results in human-derived neurons, we decided to present this method as one of the most important in vitro neurotoxicity metrics. A description of the neurotoxic effect in a quantifiable manner specific to neurones is a valuable result that can be potentially used for modelling purposes [4].

Results of the experiments reporting gene expression or concentrations of protein modifications, such as, e.g., qPCR or western blot, were excluded since there is no single, unique marker that could be traced and identified with the final endpoint, i.e., cell death.

Among the multiple experiment-specific parameters, those potentially affecting the final toxicological effect were of interest. The most often reported include the concentration of the tested

Table 1
List of columns and parameters presented in the data set.

Parameters list	Units	Description and comments
Compound	–	Name of the compound.
First author and date	–	Name of the first author and date of publication.
PMID	–	PMID of the article.
Concentration tested	µM	Each record is presented as a single concentration.
Statistical significance	binary	Cell viability might change due to various factors. This parameter tells us if it was most likely caused by applying a compound to the medium.
In vitro model	–	Name of the cells used. Only results from human-derived cells are presented.
Differentiation	binary	It is possible to differentiate some of the cell lines so that they imitate specialised cells, e.g., neurons or astrocytes. Resistance to toxic substances might be different depending on the differentiation state.
Density of cells used for plating [per well]	cells/well	The density of cells placed in the testing well represents the number of cells that compound was tested on, assuming that cells did not proliferate.
Assay endpoint	–	Measured endpoint, e.g., disruption of the cell membrane, change in metabolic activity, etc.
Drug incubation time before measurement	h	The time of incubation of cells with the drug can greatly affect the toxic effect. Usually, the longer the incubation time, the more toxic effects we observe.
Activity substrate/technique	–	Colorimetric substrate or technique that has been used to obtain a given endpoint.
Assay reactant incubation time	h	Says for how long cells were incubated with the colorimetric reactant before measurement. Usually, longer incubation times with the reactant will give a more intense colour.
Absorbance wavelength	nm	Wavelength that is assumed to be specific for the analyte. It can affect the outcome if a wavelength other than the recommended one is used.
Recording time	h	Specific for MEA records, since parameters obtained are usually the mean from that measurement time.
Average number of active electrodes per well at the time of the assay	–	Specific for MEA. Since it is difficult to make cells grow directly on the electrodes of the MEA plate, it is possible that neurons will not activate every single electrode.
Number of total available electrodes per well	–	Different models of MEA plates might have different numbers of electrodes on them.
Mean Spike Rate (mean firing rate)	% of negative control	Number of spikes that occurred in recording time divided by recording time.
Mean Burst rate	% of negative control	Number of bursts that occurred in recording time divided by recording time.
Mean network burst rate	% of negative control	Number of network bursts that occurred in recording time divided by recording time.
Inter spike interval (ISI) coefficient od Variation	% of negative control	SD of ISI divided by mean ISI informs about spike regularity. The closer to zero, the more regular the spike distribution is; if higher than one, the signal is a burst.
Burst Duration	% of negative control	Average time from the first spike in a burst to the last spike.
Number of Spikes per Burst	% of negative control	Average number of spikes in a burst.
Mean ISI within Burst	% of negative control	Mean ISI within burst.
Median ISI within Burst	% of negative control	Median ISI within burst.
Inter-Burst Interval (IBI)	% of negative control	Time interval between the last spike of a burst and the spike of a subsequent burst.
Burst Frequency	% of negative control	Number of bursts during recording time divided by recording time.
IBI Coefficient of Variation	% of negative control	Tells about burst regularity. Obtained by dividing the SD of IBI by the mean of IBI.
Burst Percentage	% of negative control	Percentage of spikes in burst when compared to all the spikes.

(*continued on next page*)

Table 1 (*continued*)

Parameters list	Units	Description and comments
Network Burst Frequency	% of negative control	Number of network bursts that occurred in recording time divided by recording time.
Network Burst Duration	% of negative control	Average time from first spike in a network burst to last spike.
Number of Spikes per Network Burst	% of negative control	The average number of spikes in the network burst.
Mean ISI within Network Burst	% of negative control	Average of mean ISIs within a network burst.
Median ISI within Network Burst	% of negative control	Average of the median ISIs within a network burst.
Number of Spikes per Network Burst/Channel	% of negative control	The average number of spikes in a network burst is divided by the number of electrodes participating in the burst.
Network Burst Percentage	% of negative control	The percentage of spikes in the network bursts when compared to all spikes.
Network IBI Coefficient of Variation	% of negative control	Measure of network burst regularity. The closer to zero, the more regular the interval is. Obtained by dividing the SD of network IBI by the mean network IBI.
Network Normalized Duration interquartile range (IQR)	% of negative control	Measure for network burst duration regularity. The larger the value, the wider the variation.
Area Under Normalized Cross-Correlation	% of negative control	The higher the value, the greater the synchronicity of the network.
Full Width at Half Height of Normalized Cross-Correlation	% of negative control	Measure for network synchronisation, the lower the value, the less synchronised the network is.
MTT	% of activity of negative control	MTT can be used to measure metabolic activity. It reacts with reducing agents, producing insoluble formazan salt crystals both intracellularly and extracellularly. It is assumed that it measures the activity of reductases, whose activity is correlated with cell viability in the culture; however, it can react to formazan by interaction with reductors present in cells and medium. Substances tested might either increase or decrease mitochondrial activity without damaging the cell. Substance can also be a reducing factor itself. Method accuracy decreases in an acidic environment [8,9].
MTS	% of activity of negative control	Improved version of the MTT assay. MTS creates water-soluble formazan, so no extra dissolution step is needed [10].
WST-8	% of activity of negative control	The same group of tests as MTT and MTS, however, WST-8 is not cell membrane permeable, so metabolic activity is measured only outside the cell. It does not require a dissolution step.
Neutral Red	% of uptake of negative control	Viable cells are able to incorporate and bind red dye in lysosomes; cells with disrupted membranes and open lysosomes lose this ability [11].
Free dsDNA	% of release when compared to negative control	The presence of free double stranded DNA is an indicator of cell death due to membrane disruption [12].
Membrane damage (LDH release)	% of negative control or% of positive control	Lactate dehydrogenase (LDH) is present inside the cell, and its activity in the extracellular liquid is a marker of cell disruption, which suggests cell death. Its activity is measured by providing a non-permeable substrate for this enzyme, which after reaction, will create a colour product [13].
Superoxide production	% of negative control	An increase in the total concentration of free radicals usually causes considerable damage to cells; so increase of any reactive oxygen species (ROS) might be the cause of cell death.
Nitrite production	% of negative control	Types of free radicals. It is possible to measure it separately from other free radicals.
Lipid peroxidation	% of negative control	Lipid peroxidation is based on measurements of the end products of lipid peroxidation. One of the types of radicals causing oxidative stress.

(*continued on next page*)

Table 1 (*continued*)

Parameters list	Units	Description and comments
Carbonylation	% of negative control	Carbonyl compounds can be a measure of increased oxidative stress since their release is caused by peroxidation.
Apoptotic cells / Living cells (Annexin V and PI staining)	Absolute% of sample or% of negative control	Annexin V marks phosphatidylserine, which is mostly present on the inner side of the cell membrane. When cells enter apoptotic cells, phosphatidylserine moves to the outer layer, which allows annexin to bind to it, releasing the fluorescent reagent (FITC). Additionally, propidium iodide present in the mixture can mark cellular DNA if the membrane is completely disrupted. This way, we can distinguish viable, apoptotic, and necrotic cells. Results of flow cytometry can be presented as an absolute value (number of cells in the sample) or as a relative value (the percent of change in apoptotic, viable, and necrotic cells when compared to the control) [14].
Alamar Blue	% of negative control	Another type of formazan salt assay. It uses the reducing capabilities of the cellular environment to create fluorescent derivatives.
Intracellular reactive oxygen species (ROS)	% of negative control or% of positive control	Measured with cell permeable reagent that, after oxidation, becomes fluorescent.
Membrane damage (trypan blue cell count)	% of dead cells	Trypan blue cannot permeate the cell membrane, thus it marks only cells that have disrupted cell membrane. Cells are counted manually by the experimenter on hemocytometer [15].
Casp-3 activity	% of negative control	Caspase 3 is an enzyme that takes part in cell apoptosis; its increased activity is a mark of cell death [16].
SD / SEM	–	Standard deviation / standard error of the mean. We collected only values presented by authors; we do not recalculate SEM to SD or backwards. In the case of MEA, all parameters can be obtained at once, and each one of them has its own SD or SEM assigned to it if available.

compound and the incubation time. All parameters reported in the database are presented in Table 1.

The dataset in the current state consists of 862 records of 61 different compounds (63 if accounting for salts). We plan to expand it in the future. All the data will be held locally as an .xlsx extension file, in a Mendeley Data repository, and as a part of the repository on the tox-portal.com website, which until recently was a website storing in-vitro cardiotoxicity study results, now expanded by the neurotoxicity data set.

4. Experimental Design, Materials and Methods

Articles were found using the PubMed search engine; only original research articles were considered. The search queries used were made of several combinations of the following phrases: “neurotoxicity”, “in-vitro”, “hiPSC”, “iPSC”, “SH-SY5Y”, “SK-N-SH”, “MEA”. The selected range of dates of publication was 2003–2023.

Data was acquired with the use of a freeware digitising tool, Web Plot Digitizer version 4.6. Plots and figures containing data were downloaded in the best available quality. If it was not possible or if the quality was unsatisfactory (after a visual check), screenshots of the highest possible resolution were taken. Because of the structure of most of the plots, the standard deviation (SD) or standard error of the mean (SEM) were obtained by subtracting the values of the mean from the values of the mean plus whiskers. Records that required correction due to problems, described in the “Limitations” section, were fixed manually. Together with concentration-response results, experimental conditions (cf. Table 1) were collected and listed in the repository.

Limitations

Since it was not possible to obtain the original raw data used in the analysed publications, all the records contain digitised data. With that, they might differ from the actual values due to limited manual digitization precision. Moreover, this way of collecting data is user-sensitive, so even if the same methodology is applied, differences between raw and digitised values might occur.

The data set presents only the critical parameters, whose modification carries the potential for a significant impact on the measured value. However, the list does not include all potentially significant parameters. Most of the analysed articles do not provide information on passage number, percent of differentiated neurons in the culture, or the final concentration of colorimetric substrate, which are parameters that have an influence on the result. Also, in the case of enzyme activity, the weight of the obtained protein for analysis is important and should be reported, yet it is not a common practice to report this parameter in neurotoxicological assessments.

When, in some cases, digitised values were approaching zero on the y-axis and it was suggested by the shape of the plot or by the authors of the original report, we assumed them to be zero. Moreover, some points on the plots were indistinguishable, so a higher discrepancy can be expected in their value. When indistinctness occurred for distribution parameters – standard deviation (SD) and standard error of the mean (SEM), they were assumed to be equal to zero, since there was no other way of evaluating them.

Ethics Statement

The authors have read and follow the ethical requirements for publication in Data in Brief and confirming that the current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

Data Availability

Dataset of in-vitro measured chemicals neurotoxicity. (Original data) (Mendeley Data).
Dataset of in-vitro measured chemicals neurotoxicity. (Original data) (Tox-portal.com).

CRediT Author Statement

Seweryn Ulaszek: Conceptualization, Data curation, Investigation, Software, Writing – original draft; **Bartek Lisowski:** Supervision, Project administration, Writing – review & editing; **Sebastian Polak:** Supervision, Project administration, Writing – review & editing.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Rohatgi Ankit, 'WebPlotDigitizer'. Pacifica, California, USA, Sep. 2022. [Online]. Available: https://automeris.io/WebPlotDigitizer
[2] S. Polak, B. Wiśniowska, A. Glinka, M. Polak, Tox-Database.net: a curated resource for data describing chemical triggered in vitro cardiac ion channels inhibition, BMC Pharmacol. Toxicol. 13 (1) (Dec. 2012) 6, doi:10.1186/2050-6511-13-6.

[3] L. Krumpholz, B. Wiśniowska, S. Polak, Open-access database of literature derived drug-related Torsade de Pointes cases, BMC Pharmacol. Toxicol. 23 (1) (Dec. 2022) 7, doi:10.1186/s40360-021-00548-0.
[4] 'U.S. EPA. 2015. ToxCast & Tox21 summary files from invitrodb_v3. Retrieved from https://www.epa.gov/chemical-research/toxicity-forecaster-toxcasttm-data on October 28, 2018. Data released October 2018.'
[5] D.L. Filer, *tcpl: toxCast Data Analysis Pipeline*. 2023. [Online]. Available: https://CRAN.R-project.org/package=tcpl
[6] T. Miyakawa, No raw data, no science: another possible source of the reproducibility crisis, Mol. Brain 13 (1) (2020) 24 s13041-020-0552–2, Dec., doi:10.1186/s13041-020-0552-2.
[7] A.C. Daly, M. Clerx, K.A. Beattie, J. Cooper, D.J. Gavaghan, G.R. Mirams, Reproducible model development in the cardiac electrophysiology Web Lab, Prog. Biophys. Mol. Biol. 139 (Nov. 2018) 3–14, doi:10.1016/j.pbiomolbio.2018.05.011.
[8] M. Ghasemi, T. Turnbull, S. Sebastian, I. Kempson, The MTT assay: utility, limitations, pitfalls, and interpretation in bulk and single-cell analysis, Int. J. Mol. Sci. 22 (23) (Nov. 2021) 12827, doi:10.3390/ijms222312827.
[9] J.G. Gomez-Gutierrez, et al., The neutral red assay can be used to evaluate cell viability during autophagy or in an acidic microenvironment in vitro, Biotech. Histochem. 96 (4) (May 2021) 302–310, doi:10.1080/10520295.2020.1802065.
[10] T.L. Riss, R.A. Moravec, A.L. Niles, et al., Assay guidance manual [Internet], Cell Viability Assays, Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD), 2013 [Online]. Available https://www.ncbi.nlm.nih.gov/books/NBK144065/ .
[11] V. Zuang, The neutral red release assay: a review, Altern. Lab. Anim. 29 (5) (Sep. 2001) 575–599, doi:10.1177/026119290102900513.
[12] A.I. Dragan, J.R. Casas-Finet, E.S. Bishop, R.J. Strouse, M.A. Schenerman, C.D. Geddes, Characterization of PicoGreen interaction with dsDNA and the origin of its fluorescence enhancement upon binding, Biophys. J. 99 (9) (Nov. 2010) 3010–3019, doi:10.1016/j.bpj.2010.09.012.
[13] P. Kumar, A. Nagarajan, P.D. Uchil, Analysis of cell viability by the lactate dehydrogenase assay, Cold Spring Harb. Protoc. 2018 (6) (2018) pdb.prot095497 Jun., doi:10.1101/pdb.prot095497.
[14] I. Vermes, C. Haanen, H. Steffens-Nakken, C. Reutellingsperger, A novel assay for apoptosis Flow cytometric detection of phosphatidylserine expression on early apoptotic cells using fluorescein labelled Annexin V, J. Immunol. Methods 184 (1) (1995) 39–51 Jul., doi:10.1016/0022-1759(95)00072-I.
[15] W. Strober, Trypan Blue Exclusion Test of Cell Viability, Curr. Protoc. Immunol. 111 (1) (2015) Nov., doi:10.1002/0471142735.ima03bs111.
[16] A.G. Porter, R.U. Jänicke, Emerging roles of caspase-3 in apoptosis, Cell Death Differ 6 (2) (1999) 99–104 Feb., doi:10.1038/sj.cdd.4400476.