Modified *Xenopus laevis* approach (R-FETAX) as an alternative test for the evaluation of foetal valproate spectrum disorder

Maria Battistoni\textsuperscript{1,2*}, Renato Bacchetta\textsuperscript{2*}, Francesca Di Renzo\textsuperscript{2}, Francesca Metruccio\textsuperscript{3}, Angelo Moretto\textsuperscript{4}, Elena Menegola\textsuperscript{2}

\textsuperscript{1}Università degli Studi di Milano, Department of Physics Aldo Pontremoli, via Celoria 16-20133 Milan, Italy. maria.battistoni@unimi.it

\textsuperscript{2}Università degli Studi di Milano, Department of Environmental Science and Policy, via Celoria 26-20133 Milan, Italy. renato.bacchetta@unimi.it, francesca.direnzo@unimi.it, elena.menegola@unimi.it

\textsuperscript{3}ICPS, ASST Fatebenefratelli Sacco, via GB Grassi 74- 20159, Milan, Italy. francesca.metruccio@unimi.it

\textsuperscript{4}Università degli Studi di Milano, Department of Biomedical and Clinical Sciences “L. Sacco”, via GB Grassi 74- 20159 Milan, Italy; present affiliation: Università degli Studi di Padova, Department of Cardiac Thoracic Vascular and Public Health Sciences, Via Giustiniani 2, 35128 Padova, Italy. angelo.moretto@unipd.it

*These authors contributed equally to this work

Corresponding author: Francesca Di Renzo, Department of Environmental Science and Policy, Università degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy. +39 02 50314804; E-mail: francesca.direnzo@unimi.it
ABSTRACT

In compliance to animal welfare 3Rs principle there is a great demand for refined tests alternative to classical mammal teratogenicity tests. We propose a refined alternative amphibian method (R-FETAX) to evaluate chemical induced embryotoxicity. The human foetal valproate spectrum disorder (FVSD) characteristics are morphological defects (including cranio-facial, neural tube defects) and behavioural alterations due to valproate (VPA) exposure in pregnancy. Vertebrate assays to evaluate FVSD include classical and alternative mammal (implying adult sacrifice), and non-mammal developmental models (zebrafish, amphibians, chick). Among these latter only zebrafish assays report in the same test both morphological and behavioural examinations. Compared to zebrafish, the amphibian *Xenopus laevis* excels having a more comparable organ development and morphology to mammalian systems. We used *X. laevis* embryos exposed during developmental specific windows to VPA therapeutic concentrations. Different VPA effects were observed depending on the exposure window: concentration-related embryo-lethal and teratogenic effects (neural tube, facial, tail defects) were observed in groups exposed at the organogenetic phylotypic stages. Neurobehavioral deficits were described using a functional swimming test at the highest VPA concentration exposure during the phylotypic stages and at any concentration during neurocognitive competent stages. Malformations were compared to those obtained in a mammalian assay (the rat post-implantation whole embryo culture method, WEC), that we used in the past to evaluate VPA teratogenicity. R-FETAX and WEC data were modelled and their relative sensitivity was calculated. We suggest the amphibian R-FETAX as a refined windowed alternative test for the evaluation of chemicals inducing both morphological and behavioural anomalies, including VPA.

KEY WORDS

R-FETAX, WEC, VPA, 3R, FVSD
1. Introduction

Since teratogenicity testing in mammals is a particular task for animal welfare, there is a great need for the development of alternative test systems. The amphibian *Xenopus laevis* and zebrafish (*Danio rerio*) embryos are becoming gold standard as embryonic vertebrate models because adult sacrifice is not necessary. The predictivity for human teratogenicity of these methods is strongly due to the strict correlation among vertebrate development at their phylotypic stages (stages where embryos of different classes express the highest degree of morphological and molecular resemblance). Valproic acid (VPA) is a widely prescribed broad-spectrum antiepileptic drug that is particularly useful for the management of generalised epilepsies [1]. The efficacy of its action and the numerous information related to the pharmacokinetics of this molecule, combined with other drugs, have extended the use of VPA as a migraine prevention and control agent and as a drug of choice in some types of schizophrenia [2]. Owing to its low adult toxicity, it is well tolerated by patients and is therefore widely used for other therapeutic purposes [3], including those for Alzheimer's and HIV diseases [4].
Finally, an innovative use of VPA has been proposed in the alcohol detoxification phase to avoid possible relapses [5].

In humans, the use of VPA, a well-known teratogen [6,7], during the first trimester of pregnancy is associated with an increased risk of major and minor malformations, including a 20-fold increase in neural tube defects (spina bifida is the most frequent) when compared to the risk in the non-exposed children population [8], cleft lip and cleft palate, cardiovascular abnormalities, genitourinary defects, developmental delay, endocrinological disorders, and axial and limb skeletal defects [9]. This wide spectrum of malformations is also known as foetal valproate syndrome (FVS) [10]. In addition, neurodevelopmental delays, including language, cognitive, and social impairments, have been reported in infants exposed in utero to VPA, and a complete picture of morphological and functional impairments is described as Foetal Valproate Spectrum Disorder (FVSD) [11]. Epidemiological studies have linked VPA therapeutic maternal plasma concentrations (280-700 µM, for a daily dose of 300-2000 mg) to higher risks of autism spectrum disorders and attention-deficit/hyperactivity disorder [12–15]. Moreover, VPA levels in cord serum are often higher than those in the mother and may be up to five times higher than the levels in maternal serum at term [13,16].

Due to this plurality of effects, different animal models have been used in the past to mimic the teratogenic effects of VPA on human embryos to elucidate the mechanism of its teratogenic action. Developmental toxicity tests showed that VPA caused dose-related teratogenic effects in all investigated species. Malformations of multiple organs have been described after dosing with 150-800 mg/kg/day during organogenetic period, including developmental anomalies of the nervous system (mainly exencephaly) in mice and hamsters [17–20] and craniofacial and axial malformations in non-human primates, mice, rats, and rabbits [21–26]. Teratogenic dose of 400 mg/kg correspond in rats to a plasmatic concentration of about 2300 µM after 1 hr [27], while in mouse the plasmatic concentration is of about 3600 µM after 30 min [28]. Rodents were the most used animal model; in particular, mice were demonstrated to be the most susceptible to the induction of neural tube defects,
although with significant differences depending on the strain used [29]. In rats, severe delays in the neural tube closure have been documented together with axial malformations [30,31]. VPA-related behavioural disorders were investigated in rodents exposed in utero, showing characteristics similar to those of the human condition (reduction in the number of social explorations, repetitive digging behaviour, and increase in locomotor repetitive/stereotyped activity) [32].

In compliance with the 3Rs (Replacement, Reduction, Refinement) principle [33], different alternative models have been evaluated for teratological purposes. The post-implantation rodent whole embryo culture (WEC) approach was proposed by New in 1978 [34], validated [35] and then applied to better control experimental conditions, directly expose embryos without animal treatment, and reduce the number of adult animals and embryos needed for each test (Refinement and Reduction, but pregnant dams sacrifice is still needed). In WEC, VPA exposure (from 750 to 1500 µM, comparable to the concentrations used in our present work) resulted in brain and branchial defects, irregular dorsal midline, and changes in transcriptomics [36–38]. Recently, our research group reported dose-related teratogenic effects of VPA (from 31.25 to 375 µM) in a WEC model with branchial and extra-branchial (neural tube and axial) defects [39]. However, WEC test does not provide any information on behaviour and/or neurological deficits.

Non-mammalian vertebrate models have been proposed instead of mammal models to use lower vertebrate. These models showed that i) in the amphibian X. laevis 1000 - 20000 µM VPA exposure (concentrations exceedingly above the human therapeutic concentrations more than 5 fold) induced severe effects in different body districts (tadpoles with severe abnormalities in different body districts and severe developmental delays) [40,41]; ii) in zebrafish exposed to therapeutic concentrations (from 5 to 1500 µM, in line with the concentrations used in our present work) VPA induced disruption of heart looping, haematopoiesis, cranio-facial development, liver and pancreas development, oedema and brain deformities, shortening, folding of the tail, small eyes [41–45]; and iii) in chicken embryos cultured in vitro, VPA (10 µl/embryo of a 300mM solution) interfered with the somitogenesis process.
[46]; chicken embryos exposed to VPA (2–8 mg/kg for a 50 g egg, considered from the Authors close to the range of possible human exposure) in ovo showed increased mortality, growth delay, neural tube, cardiovascular, cranio-facial, limb, and skeletal anomalies [47].

Recently, zebrafish exposed to VPA during the whole embryogenic period [43] and X. laevis exposed to VPA at late tadpole stages [48] were also proposed as models for behavioural studies on the VPA-related autism spectrum disorder: in zebrafish hyperactivity of embryo/larvae movement behaviours and increase of larval social behaviours, in X. laevis abnormal visual avoidance and schooling behaviours. Results from both studies, were consistent with the observations made in rodent models and with human characteristics of autism and autism-related disorders, suggesting that VPA exposure affects the main developmental process that is evolutionarily conserved across species. Considering X. laevis as an evolutionary closer model in respect to the teleost fish zebrafish [49], this study aimed to evaluate the applicability of a modification of frog embryo teratogenesis assay: Xenopus (FETAX) termed refined FETAX (R-FETAX) for multiple aspects of FVSD panel including both some morphological and behavioural effects (craniofacial and neural tube defects and neural functional impairment). Therefore, the standard FETAX exposure period (according to ASTM standard guide E1439 [50], from mid-blastula to tadpole) was subdivided into different exposure windows for different targets: i) pre-organogenetic period, ii) organogenetic period, and iii) spontaneous swimming acquisition period. VPA teratogenic effects were evaluated by exposing the embryos during the phylotypic organogenetic period and compared to those obtained in an in vitro mammalian model (WEC) using PROAST software analysis. In addition to the classical teratological evaluation based on morphological observations, functional deglutition and neurobehavioral tests were performed.

2. Material and Method
2.1. R-FETAX methodology

Amphibian *X. laevis* adults (Nasco, USA) were maintained in an automatic breeding system (TecnoPlus, Techniplast, Italy) under controlled water conditions (T = 20±2°C; pH = 7.5±0.5; conductivity = 1000±100 μS), 12-h light/dark cycle (light from 7:00 AM to 7:00 PM) and fed with a semisynthetic diet twice a week (XE40 by Mucedola; Settimo Milanese, Italy). In compliance with the refinement of the 3R principle [33] and in contrast to the classic FETAX methodology (ASTM standard guide E1439), embryos were obtained without human chorionic gonadotropin injection from overnight natural mating in a mating system with controlled humidity and air/water temperature. The collected embryos were cleaned by gentle swirling in a 2.25% L-cysteine solution with an arranged pH of 8.0 and rinsed several times in FETAX solution, whose composition was 625 mg/L NaCl, 96 mg/L NaHCO₃, 30 mg/L KCl, 15 mg/L CaCl₂, 60 mg/L CaSO₄·2H₂O, and 70 mg/L MgSO₄. Normally cleaved embryos at the mid-blastula stage (stage 8, according to Nieuwkoop and Faber [NF] [51]) were selected for testing. During the whole test time (6 days, considering 0 the morning after egg deposition), samples were maintained in a thermostatically controlled FETAX solution (5 embryos per Petri dish) at 23°C from NF stage 8 until day 6, corresponding to NF stage 46, as evaluated in preliminary tests on unexposed larvae. Considering NF 47, the last not independently feeding larval form [52] and excluded in the list by EU directive 2010/63/EU on the protection of animals used for scientific purposes [53], R-FETAX was limited to 6 days. Exposures covered the whole length of the standard FETAX procedure (NF stage 8-46) or was limited to windows covering some developmental phases considered of interest, such as: *i*) from day 0 to day 0.5 (NF stage 8-13, pre-organogenetic period); *ii*) from day 0.5 to day 2 (NF stage 13-37, organogenetic period); *iii*) from day 2 to day 6 (NF stage 37-46, spontaneous swimming acquisition period); *iv*) from day 0.5 to day 1 (NF stage 13-26, early neurula-early tailbud, representing the phylotypic stage and therefore the vertebrate common window useful for teratogenesis purposes) (Fig. 1).
2.1.1 Embryolethality evaluation

The first set of tests was performed to determine embryolethality. VPA (sodium salt, Sigma) was dissolved in FETAX solution to obtain final concentrations of 0–500–750 and 1500 µM VPA based on a range-finding test. Different exposure windows were tested (Fig. 1). Embryolethality was evaluated throughout the test, and dead samples were removed from the experimental groups. At the end of the test, living larvae were overdosed at 4°C with an anaesthetic (0.5% MS222, Sigma, dissolved in FETAX solution). Euthanized larvae were rinsed in FETAX, fixed in ethanol 50% and conserved in ethanol 70%.

2.1.2 Teratogenicity evaluation

Teratogenicity was evaluated in embryos exposed during the species-agnostic organogenetic period (NF stages 13-26). Embryos were exposed to 0–500–750–1500 µM VPA. On day 6, tadpoles were subjected to the new deglutition test, freshly evaluated for morphological alterations, fixed, and processed for cartilage and muscle double staining. The deglutition test was designed considering a recent study reporting that X. laevis larvae can ingest polystyrene plastic microparticles visible in the intestine during external evaluation [54]. Larvae were maintained for 30 min at 23°C in FETAX solution containing 12.5 µg/mL red polystyrene microparticles (1µm diameter, Sigma), euthanized, and fixed as described above. Larvae were observed under a stereomicroscope (Leica) for morphological evaluation and red staining (deglutition test positive) at the level of the gut: only larvae with functioning jaws can ingest microplastic.

Thereafter, tadpoles were processed for double staining of tissues rich in mucins (alcian blue, staining cartilage and connectives) or rich in calcium (alizarin red, staining vessels, muscular, and nervous tissues). Staining was performed according to the partially modified version of the method described by Walker and Kimmel [55]. All reagents were purchased from Sigma, Italy. Briefly, the alizarin solution (0.5% alizarin red in distilled water) was mixed 1:5 to ethanol 70% (red staining) and then
dissolved 1:100 in a blue staining solution (0.02% alcian blue in ethanol 70% containing 40mM MgCl₂). Samples were incubated at RT under stirring overnight in the double staining solution, rinsed in ethanol 70%, and observed under a dissecting microscope (Leica). Cartilages appeared dark blue, connectives light blue, nervous system, and muscles light red. The flat mount technique [56] was applied to evaluate the mouth articular regions in detail.

2.1.3 Neuro-behavioural evaluation

The Neurobehavioral evaluation was performed on embryos exposed to 0-500-750-1500 µM VPA during the phylotypic stages (NF stages 13-26) or spontaneous swimming acquisition period (NF stages 37-46). According to Currie et al. [57], NF stages 37-46 represent the transition from dormant life to progressive free-swimming locomotion at the onset of active feeding. The motor behaviour of the tadpoles was evaluated on day 6 by setting a swimming test developed by modifying the method described by De Felice et al. [54]. Tadpoles were transferred into a 27 mm plastic cylinder, representing the arena, placed inside a 90 mm plastic Petri dish filled with FETAX solution on an under-illuminated stereomicroscope (Leica). Larvae were allowed to acclimate for 1 min before recording. Videos were taken from above for 30 s using a 1080p HD 30 fps digital camera and subsequently analysed using the AnimalTracker plugin [58] and the free image processing program ImageJ [59]. The protocol was set considering the normal swimming activity of a group of unexposed larvae (swimming close to the edge of the arena, outer ring). The arena diameter was set at 1, and the evaluation was performed on the activity of each larva in the inner circle (with a diameter 0.75 of the arena diameter) and in the outer ring (0.25 of the Arena) (Fig. 2A). The total immobility time, total distance (mm), distance (mm) in the outer ring and the inner circle, swimming speed (distance/total time-immobility time), time spent in the inner circle/total time, and the distance inner circle/total distance (%) were considered as swimming activity endpoints. After the test, larvae were transferred to fresh FETAX, euthanized, fixed, and conserved as described above.
2.2. Rat WEC

Virgin female Crl:CD rats (Charles River, Calco, Italy), housed in a thermostatically maintained room (T=22±2°C, relative humidity = 55±5%) with a 12-h light cycle (light from 6:00AM to 6:00 PM), free access to food (4RF21, Charles River, Calco, Italy), and tap water were caged overnight with males of proven fertility. All animal experimental protocols were approved by the Ministry of Health, Department of Veterinary Public Health, Nutrition and Food Safety Committee. The animals were treated humanely and with regard to the alleviation of suffering. Embryos were explanted in sterile Tyrode solution (Sigma) from untreated pregnant rats at E9.5 early neurula stage, 1–3 somites; day of positive vaginal smear = 0) and cultured according to the method proposed by New [34] in 20-mL glass bottles (five embryos/bottle), containing 4 mL of culture medium composed of undiluted heat-inactivated rat serum supplemented with antibiotics (penicillin 100 IU/mL culture medium and streptomycin 100 µg/mL culture medium, Sigma). The culture was performed in triplicate for each group. The bottles, inserted in a thermostatic (37.8°C) roller (30 rpm) apparatus, were periodically gas equilibrated according to Giavini et al. [60]. Tested concentrations of VPA (sodium salt, Sigma) were dissolved in Tyrode solution and added (80µl/bottle) to the culture medium of the treated group to obtain final concentrations of 0–250–500–750 µM. VPA concentrations were selected based on Metruccio et al. [39] to obtain marked branchial and extra-branchial defects, considering the human therapeutic plasma concentration range (280-700 µM) and the concentration used in other in vitro models [21,30,29,43]. After 48 h of culture, embryos were morphologically examined under a dissecting microscope (Leica) to evaluate any embryonic abnormalities.

2.3. Statistical analysis and mathematical modelling

Quantal data, expressed as percentages, were analysed using the Chi-square test. Continuous data, expressed as mean and standard deviation, were analysed using Student’s t-test. The level of
significance was set at $p < 0.05$. The benchmark dose (BMD) approach was applied using PROAST (67 version), a software package developed by the Dutch National Institute for Public Health and the Environment (RIVM) (www.proast.nl) for the statistical analysis of dose-response toxicological data. An important feature of PROAST is that it allows comparing dose responses among subgroups (in our case, WEC and R-FETAX data). Data were modelled to characterise the single dose-response curves on WEC or FETAX samples with the brain, facial, and axial abnormalities, setting the benchmark dose (BMD) at $25–50–75\%$ benchmark response (BMR). After modelling the results obtained for each parameter with WEC and R-FETAX, the log-likelihood ratio test was applied to assess the equal steepness assumption. When the test passed, the relative potency factors (RPFs) of R-FETAX versus WEC were derived to calculate the relative sensitivity of R-FETAX compared to WEC. The exponential model family equations were selected to describe the dose-response curves and obtain the RPFs versus WEC.
3. Results

3.1. R-FETAX

3.1.1. Evaluation of embryolethality and selection of exposure windows

In the first set of experiments, a clear concentration-dependent embryolethal effect was observed in groups exposed to VPA during the standard FETAX exposure period NF stages 8-46 and during the organogenetic window NF stages 13-37 (Table 1). Lethality mainly occurred on day 4 (NF stage 42-43). These data suggest that VPA causes impairment of vital organ functionality due to irreversible effects on the organogenetic program. No significant lethal effects of VPA exposure were observed in the groups exposed to VPA during NF stages 8-13, 37-46. Limiting the exposure to the species-agnostic window (NF stages 13-26), only the highest VPA concentration leads to a significant embryolethal effect. Because NF stages 13-26 correspond to the phylotypic stages and rat developmental stages in WEC, this window of exposure was selected for the teratogenicity evaluation. Neurobehavioral tests were performed on tadpoles exposed to VPA during NF 13-26 or NF 37-46.

3.1.2. Teratogenicity evaluation

In terms of morphological evaluation, normal tadpoles displayed well-distinguishable craniofacial (circular or oval oral opening with prominent lower jaw, linear brain parallel to the body axis, and well-expanded gill basket), abdominal (with properly twisted intestine and pronephros visible), and caudal (linear notochord, regularly organised axial muscles) portions (Fig. 3A). Some controls showed a reduced lower jaw or bent tail (Table 2). The exposure to VPA during NF stages 13-26 was teratogenic (Table 2), inducing a concentration-dependent manner, brain abnormalities (swollen brain), and facial defects (abnormal oral opening with reduced jaws) (Fig. 3B). The deglutition test showed a significant trend in tadpoles unable to ingest (Table 2), suggesting an impairment in deglutition capability due to severe mouth malformations. Mouth abnormalities were further
evaluated after double staining of muscles and cartilages. Visible articular spaces were observed in larvae able to ingest microparticles (Fig. 3 C, E, G), whereas larvae exposed to VPA negative for deglutition were characterised by fused cartilaginous facial elements with a consequent deviation of the correlated muscles (Fig. 3 D, F, H). Despite a significant trend for tail abnormalities (bent), no structural alterations were observed at the level of the axial muscles. No cardiovascular abnormalities were observed.

3.1.3. Neuro-behavioural evaluation

In the control group, tadpoles spent most of their time in the outer ring, swimming in a circular regular manner (Fig. 2B). In contrast, tadpoles exposed to VPA 1500 µM during the teratological window (NF 13-26) showed an abnormal swimming route that spent more time in the inner circle (Table 3, Fig. 2C). Neurotoxicological effects were observed in tadpoles exposed late to VPA 500-1500 µM (NF stages 37-46) with reduced speed, causing a reduced covered distance (Table 3).

3.2. Rat WEC

At the end of 48 h of culture, normal embryos reached the phylotypic stage and were characterised by a tripartite encephalon (forebrain, midbrain, hindbrain) with enlarged ventricles, three well-separated branchial arches, and well-visible metameric pairs of somites in the dorsal axial region (Fig. 4A). VPA was teratogenic at any tested concentration, inducing specific abnormalities in different body districts with significant linear dose-response trends (Table 4). The specific target structures were the brain (swollen), facial primordia (branchial arches, fused, and/or reduced), axial structures (somites fused, and/or somites misaligned), and no limb or cardiovascular abnormalities were observed (Fig. 4B).

3.3. Comparison of teratogenic effects observed in WEC and R-FETAX
PROAST analysis was performed to compare teratogenic data recorded in the two experimental tests (WEC and R-FETAX) grouped as brain, facial, and axial outcomes. These districts were the targets of VPA in both experimental models.

3.3.1. **BMD inducing 25-50-75 % brain, facial and axial abnormalities**

Data were modelled to characterise the single dose-response curves on WEC or R-FETAX samples by setting the BMDs at 25–50–75% BMR (Table 5). Data on tail defects observed with R-FETAX affected only a marginal percentage of tadpoles and were not accepted by any model considered by PROAST.

3.3.2. **Comparison between brain and facial results from WEC and R-FETAX**

The first step consisted of comparing the fits to the single datasets with the fit to the combined dataset (Table 5), using the exponential model in both cases. The log-likelihood ratio test showed that the equal steepness assumption was not rejected (p = 0.87 for the brain; p=0.57 for facial) (Table 5). As the steepness result was homogenous, the RPFs were estimated using the combined model fit (Table 5). Concerning brain effects, BMD CIs and RPF CIs suggest that the two experimental tests can be considered similar in sensitivity. For facial abnormalities, even if R-FETAX potency factor resulted somehow lower than WEC (CI not including 1), the value is however of the same order of magnitude (Table 5).
4. Discussion

This study aimed to evaluate the applicability of an amphibian method (R-FETAX) alternative to mammal protocols in the study of a panel of features of the FVSD.

The amphibian *X. laevis* is an alternative animal model suitable for addressing developmental questions, toxic effects, and human disease mechanisms [61]. Exposure of *X. laevis* embryos during morphogenetic periods has been previously used to evaluate VPA-related teratogenesis [40,62] or evaluate toxicological patterns [62,41]. In contrast to our exposure protocol, these studies were based on pulse exposure to extremely high VPA concentrations (5000-20000 µM), inducing severe phenotypes (plurimalformed tadpoles with severe abnormalities in different body districts and with severe developmental delays). VPA exposure at late tadpole stages (NF stages 42-47/49, including for some tests independently feeding larval forms, included in the EU directive 2010/63/EU [53]) was associated with behavioural effects [48]. The proposed R-FETAX applies a refinement on adult animal utilization, using early embryos from naturally crossed *X. laevis* without the need to use embryos from hormone induced adults (a regulated procedure). In a six-day experiment (covering stages not reaching the legal regulated timepoint of independent feeding) this method allows to evaluate different morphological and functional outcomes, depending on the specific selected exposure window (during pre-organogenetic, organogenetic, and functional differentiation periods).

For teratological purposes, the NF stages 13-26 represent the specific window at which all vertebrates develop with high similarity at morphological and molecular levels (phylotypic window). Our data indicate that exposure to the phylotypic window induced specific teratogenic effects (brain and facial abnormalities) in the amphibian model (R-FETAX) at concentrations realistic for human conceptuses (considering a cord serum concentration up to 5-fold higher than the 280-700 µM plasma maternal concentration detected after therapeutic doses).

In the present work results obtained using amphibian model (R-FETAX) were compared with those obtained in the mammalian model (WEC) used in the past to evaluate teratogenic activity of different
The effects of R-FETAX and WEC on the brain were similar (dilated ventricles), and the relative sensitivities of the two methods (as evaluated by PROAST) were comparable. The effects on the brain observed in WEC and R-FETAX tests are consistent with the NTD reported in different animal models, such as exencephaly [63,19,20,18] and with human VPA-related brain defects [13]. The swimming test performed in R-FETAX also correlated VPA exposure during the teratogenic window with behavioural alterations (abnormal swimming route). This effect is not unexpected and is consistent with the cognitive impairments reported in children with FVS [64]. This effect, observed only at the highest concentration level, could be related to malformative patterns or to a specific neurodegenerative effect at the time of exposure. The evaluation of cell death at the end of VPA exposure could be interesting in further investigation.

We applied the swimming test to assess behavioural disorders also in larvae exposed during NF stages 37-46 (behavioural window) based on the evidence described by James et al. [48] reporting neurodevelopmental deficits after VPA exposure during late tadpole stages (NF 42-47/49). Behavioural assessment of the development of low vertebrates is becoming popular. Behavioural disturbances determine the organism’s fit and survival failure [65] and have been reported with or without gross morphological anomalies. Therefore, behavioural tests are considered sensitive and important endpoints for toxicological studies [66]. Validated and not yet validated behavioural tests have been reported in the literature for both zebrafish and *Xenopus* [67,68,54,69]. Our swimming test was designed to evaluate behavioural abnormalities at NF stage 46, considering both route and speed variations, and showed behavioural deficits (reduced swimming acquisition period (NF stages 37-46) [57] at any VPA concentration. This finding is consistent with the most recent evidence reporting VPA-related neurodevelopmental deficits in children not showing the typical morphological facies of FVS [11].

Besides neural tube defects and neurological impairment, among FVS abnormalities, our tests allowed to evaluate only craniofacial and axial defects. For the abnormalities at the facial elements,
different endpoints were evaluated in WEC and R-FETAX: WEC embryos were evaluated for branchial arches (the facial primordia resulted fused and/or reduced), R-FETAX tadpoles were evaluated for elements derived from branchial arches (quadrate, infrarostral, and Meckel’s cartilages derived from the first branchial arch; ceratohyal and basihyal derived from the second branchial arch) and for a functional deglutition test. The effects of VPA on branchial structures have been previously reported using WEC at concentrations comparable to those used in the present work [36,37,39] and facial abnormalities are known to be characteristic of FVS in both humans [13] and several other animal models [21,24,25].

For facial abnormalities evaluation, R-FETAX sensitivity was of the same order of magnitude of WEC. The proposed deglutition test resulted in a rapid method for detecting fresh samples of facial abnormalities inhibiting deglutition functionality (fusion among articular elements). This new method, not appliable in WEC assay, is proposed for the rapid screening of chemicals with facial morphogenesis as a potential target, giving information on the functional endpoint of severe morphological defects.

Axial vertebral defects have been reported in rodents and rabbits exposed to VPA in vitro (WEC, somites fused, misaligned, abnormal) or in utero (vertebral/ribs fused, duplicated, misaligned, with homeotic transformations) [21,23,26,70,71,2,39]. VPA interferes with somitogenesis in chicken embryos [46]. Notably, abnormalities of the axial skeleton have not been described in human FVS and have not been reported in the alternative animal models of X. laevis and zebrafish. In X. laevis, at the tadpole stages, the axial structures are supported only by the notochord and vertebrae are absent until NF stage 54 (www.xenbase.org). In an alternative to the evaluation of vertebral structures, in the present work other somite derivatives (axial muscles) were evaluated and were not affected by VPA exposure. In summary, the cases of bent tails could not be connected to the axial defects observed in the WEC. In an in ovo chicken model, topical VPA exposure also induced limb and vascular anomalies [47], which were not detected in our models. One of the disadvantages of FETAX
and WEC is that limb defects and cardiac septal defects cannot be properly detected (FETAX: larvae will develop limbs later in development, at the time of metamorphosis; WEC: rat embryos only show limb buds, in which only eventual developmental delay can be detected; in both models, the heart is not totally septed yet). The proposed method was unable to valuate other aspects of FVSD such as cardiac and limb defects. To assess heart issues R-FETAX could be a good model, because early heart developmental processes are highly conserved between amphibians and mammals. In addition, heart defects are not necessarily lethal in *Xenopus* because the early development is independent from a functional circulatory system [72]. However, specific not rutinarily techniques (whole-mount in situ hybridization, immunohistochemistry with anti-tropomyosin antibody, scanning electron microscopy, transmission electron microscopy, optical coherence tomography) are required to visualize and detect improperly looped hearts, failed chamber formation, and abnormal connective structure [73]. In order to evaluate limb malformations, by contrast, our proposed assay should be enriched, adding an exposure window included in the EU directive on the Protection of Animals Used for Scientific Purposes [53]. Similarly to what proposed by Fort and colleagues [74], the assay should go on till stage NF 54 (30 days after fertilization).

Our windowed exposure of *X. laevis* embryos, the detailed morphological evaluation, with the addition of deglutition and swimming tests, allows the evaluation of limited aspects of FVSD panel but is a rapid, unexpensive and not under EU legal constraints test. We suggest R-FETAX protocol as a first screening test to identify substances warrants a more in-depth evaluation and then to use an enriched protocol to have a more complete picture of the whole VPA-like syndromic aspects. Finally, it is relevant to underline that the concentrations used in this study cover the range found in human plasma, and our exposure windows are devoted to evaluating specific developmental periods crucial during human development (phylotypic stage to assess the teratogenesis assay; neurobehavioral sensitive window to assess neurotoxicity).
Literature and our present results support the hypothesis that VPA affects early developmental processes that are evolutionarily conserved across species, encouraging the use of alternative models for studying VPA related effects in pregnancy.

In the past, WEC was considered the best alternative to in utero exposure for assessing a wide range of teratogenic agents, including drugs, pesticides, environmental chemicals, nutritional excesses and deficiencies, and physical factors. However, in recent decades, the use of low vertebrate animal models has been encouraged as an alternative to mammals in compliance with the 3R principles [33]. The results of the present study indicate that *X laevis* embryos are an adequate alternative to mammals for the study of different aspects of FVSD, showing both morphological and neurological effects at concentrations covering also those reported in human maternal plasma [12,14]. It should be noted that the use of not hormonal injected adults and not feeding larval stages is not subject to EU mandatory rules or regulations, suggesting R-FETAX as an alternative model for developmental toxicity purposes without the need of authority permission [75–77].

Similarly to the classical FETAX, the selection of windows of exposure depending on the expected effect, minimises embryolethal effects and allows the evaluation of complex pictures of diseases. This could be useful for rapidly testing several valproate analogues or other neuroactive substances with suspected similar effects and, possibly, the mode of action.

In conclusion, R-FETAX may be a promising alternative test to evaluate a spectrum of complex developmental disorders induced by chemicals, including VPA and its analogues. We suggest that its use as a screening method to identify substances warrants a more in-depth evaluation.
5. References


**Figure 1: R-FETAX protocol: exposure windows.** Exposures covered the whole length of the standard FETAX procedure (NF stages 8-46) or was limited to specific windows: *i*) NF stage 8-13, pre-organogenetic period; *ii*) NF stage 13-37, organogenetic period; *iii*) NF stage 37-46, spontaneous swimming acquisition period; *iv*) NF stage 13-26, early neurula-early tailbud, representing the phylotypic stage and therefore the window for species-agnostic teratogenesis purposes.

NF = stages determined according to Nieuwkoop and Faber, 1956.
**Figure 2: Swimming test.** A) Arena with the Inner circle (with diameter 0.75 of the arena diameter) and Outer ring (0.25 of the arena) set up for the evaluation of swimming parameters; the yellow line represents the swimming route; the black arrow (→) is a tadpole. B-D) Examples of swimming route: control (B); tadpole exposed to 1500 µM VPA at NF stages 13-26 (C); and tadpole exposed to 1500 µM VPA at NF stages 37-46 (D). Note the irregular route in C and the reduced distance swum in D.
Figure 3: Phenotypes observed at the end of R-FETAX after exposure during the phylotypic window (NF 13-26). A-B) morphological appearance of larvae at fresh evaluation. A) Normal tadpole displaying oral opening with prominent lower jaw (lj), linear brain (b) parallel to the body axis, intestine with ingested red microparticles (g), and regularly organized tail muscles (tm). B) Tadpole exposed to VPA with the swollen brain (#) and abnormal oral opening with reduced jaws (→). Note the absence of red stain in the intestine.

C-D) Appearance of A-B samples after whole-mount double staining. C) jaw elements are well visible at the lateral view: Meckel’s cartilage (m), quadrate cartilage (q); ceratohyal cartilage (c).

E-H) Detailed evaluation of muscles and cartilages after flat-mount. E, G) flat-mount of facial elements of A/C (positive for deglutition) showing well visible articular spaces (arrowhead) in the larva able to ingest microparticles; infrarostral (i), Meckel’s cartilage (m), quadrate cartilage (q); ceratohyal cartilage (c), basihyal cartilage (b); orbitohyoideus muscle (*). F, H) flat-mount of facial elements of B/D (negative for deglutition) showing fusion of the different facial elements (white asterisks) with a consequent deviation of the correlated muscles (*).
Figure 4: Phenotypes observed in WEC. A) Normal phenotype characterized by tripartite encephalon: forebrain (°), midbrain (§), hindbrain (#) with enlarged ventricles, three well-separated branchial arches (square) and well visible metameric pairs of somites (*) on the dorsal axial region. B) Specific abnormal phenotype observed in VPA exposed embryos: swollen brain (@), fused and reduced branchial arches (square), fused and misaligned somites (>).
Table 1: Embryolethality (%) in controls and in groups exposed to VPA in different exposure windows - R-FETAX. **significant linear trend with p<0.01 (chi-square test).

<table>
<thead>
<tr>
<th>Exposure window</th>
<th>CONTROL</th>
<th>VPA 500 µM</th>
<th>VPA 750 µM</th>
<th>VPA 1500 µM</th>
</tr>
</thead>
<tbody>
<tr>
<td>NF st 8-46</td>
<td>4</td>
<td>68</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>(N= 166)</td>
<td>(N= 178)</td>
<td>(N= 117)</td>
<td>(N= 25)</td>
<td></td>
</tr>
<tr>
<td>NF st 8-13</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>(N= 60)</td>
<td>(N= 58)</td>
<td>(N= 59)</td>
<td>(N= 38)</td>
<td></td>
</tr>
<tr>
<td>NF st 13-37</td>
<td>9</td>
<td>64</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(N=23)</td>
<td>(N=25)</td>
<td>(N=26)</td>
<td>(N=25)</td>
<td></td>
</tr>
<tr>
<td>NF st 37-46</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(N= 24)</td>
<td>(N= 30)</td>
<td>(N= 30)</td>
<td>(N= 20)</td>
<td></td>
</tr>
<tr>
<td>NF st 13-26</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>(N= 121)</td>
<td>(N= 111)</td>
<td>(N= 94)</td>
<td>(N= 73)</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Teratogenic effects (%) in controls and in groups exposed to VPA at NFst 13-26 - R-FETAX. **significant linear trend with p<0.01 (chi-square test); NFst= Nieuwkoop and Faber stages

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (N= 25)</th>
<th>VPA 500 µM (N= 25)</th>
<th>VPA 750 µM (N= 24)</th>
<th>VPA 1500 µM (N= 23)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRAIN</td>
<td>Swollen</td>
<td>0</td>
<td>26</td>
<td>64</td>
</tr>
<tr>
<td>FACIAL STRUCTURES</td>
<td>Abnormal</td>
<td>12</td>
<td>28</td>
<td>58</td>
</tr>
<tr>
<td>TAIL</td>
<td>Bent</td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Deglutition test: negative</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>60</td>
</tr>
</tbody>
</table>
Table 3: Swimming parameters (M±SD) evaluated in controls and in groups exposed to VPA-R-FETAX. a p<0.05 vs. CONTROL group; aa p<0.01 vs. CONTROL group; b p<0.05 vs. VPA 500; c p<0.05 vs. VPA 750.

<table>
<thead>
<tr>
<th></th>
<th>Immobility Time (sec)</th>
<th>Distance (mm)-SUM</th>
<th>Speed (mm/sec) without immobility time</th>
<th>Time in the inner circle/tot time (%)</th>
<th>Distance Inner Circle/ Total Distance (mm/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CONTROL</strong> (N= 18)</td>
<td>5.36 ± 8.94</td>
<td>335.24 ± 191.45</td>
<td>13.05 ± 5.86</td>
<td>0.02 ± 0.05</td>
<td>0.03 ± 0.05</td>
</tr>
<tr>
<td><strong>VPA 500 st 12-26</strong> (N= 14)</td>
<td>14.49 ± 10.55</td>
<td>190.41 ± 153.92</td>
<td>a 12.17 ± 3.50</td>
<td>0.05 ± 0.10</td>
<td>0.07 ± 0.11</td>
</tr>
<tr>
<td><strong>VPA 750 st 12-26</strong> (N= 14)</td>
<td>8.05 ± 8.78</td>
<td>220.10 ± 96.80</td>
<td>10.35 ± 2.58</td>
<td>0.05 ± 0.08</td>
<td>0.07 ± 0.10</td>
</tr>
<tr>
<td><strong>VPA 1500 st 12-26</strong> (N= 15)</td>
<td>10.27 ± 9.17</td>
<td>201.06 ± 107.32</td>
<td>10.43 ± 2.96</td>
<td>0.23 ± 0.31</td>
<td>0.22 ± 0.20</td>
</tr>
<tr>
<td><strong>VPA 500 st 37-46</strong> (N= 15)</td>
<td>8.48 ± 10.59</td>
<td>175.58 ± 117.93</td>
<td>a aa 7.70 ± 2.80</td>
<td>0.01 ± 0.03</td>
<td>0.07 ± 0.14</td>
</tr>
<tr>
<td><strong>VPA 750 st 37-46</strong> (N= 15)</td>
<td>8.58 ± 10.59</td>
<td>162.80 ± 109.82</td>
<td>a aa 7.27 ± 2.51</td>
<td>0.08 ± 0.20</td>
<td>0.12 ± 0.22</td>
</tr>
<tr>
<td><strong>VPA 1500 st 37-46</strong> (N= 15)</td>
<td>14.68 ± 12.48</td>
<td>96.48 ± 83.12</td>
<td>5.94 ± 1.84</td>
<td>0.13 ± 0.32</td>
<td>0.15 ± 0.32</td>
</tr>
</tbody>
</table>

Table 4: Teratogenic effects (%) in controls and in groups exposed to VPA-WEC. **significant linear trend with p<0.01 (chi-square test).

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (N= 13)</th>
<th>VPA 250 µM (N= 20)</th>
<th>VPA 500 µM (N= 12)</th>
<th>VPA 750 µM (N= 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BRAIN</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swollen</td>
<td>8</td>
<td>50</td>
<td>25</td>
<td>100</td>
</tr>
<tr>
<td><strong>BRANCHIAL ARCHES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tot abnormal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fused</td>
<td>0</td>
<td>30</td>
<td>67</td>
<td>100</td>
</tr>
<tr>
<td>reduced</td>
<td>0</td>
<td>25</td>
<td>17</td>
<td>100</td>
</tr>
<tr>
<td><strong>SOMITES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tot abnormal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fused</td>
<td>0</td>
<td>15</td>
<td>17</td>
<td>100</td>
</tr>
<tr>
<td>misaligned</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 5: Parameters obtained by PROAST analysis, fitting separate datasets for the two experimental tests and combined datasets. Benchmark dose (BMD) for benchmark responses (BMR) at 25-50-75%. Overlapping confidence intervals (CI) obtained in the two tests are labelled in grey. Relative potency factors (RPF).

<table>
<thead>
<tr>
<th></th>
<th>Brain</th>
<th>Facial</th>
<th>Axial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WEC</td>
<td>R-FETAX</td>
<td>WEC</td>
</tr>
<tr>
<td><strong>BMD for BMR 25% (CI)</strong></td>
<td>223 (150-572)</td>
<td>579 (491-618)</td>
<td>241 (147-349)</td>
</tr>
<tr>
<td><strong>BMD for BMR 50% (CI)</strong></td>
<td>383 (331-609)</td>
<td>644 (583-688)</td>
<td>362 (281-448)</td>
</tr>
<tr>
<td><strong>BMD for BMR 75% (CI)</strong></td>
<td>689 (516-3130)</td>
<td>735 (694-791)</td>
<td>499 (409-644)</td>
</tr>
<tr>
<td><strong>loglikelihood ratio test (p)</strong></td>
<td>0.81</td>
<td>0.15</td>
<td>nd</td>
</tr>
<tr>
<td><strong>RPF (index: WEC) (CI)</strong></td>
<td>1 (0.75-1.01)</td>
<td>0.87 (0.45-0.72)</td>
<td>1 (0.45-0.72)</td>
</tr>
</tbody>
</table>