Relationship between increasing concentrations of two carcinogens and statistical image descriptors of *foci* morphology in the Cell Transformation Assay

Journal:	*Journal of Applied Toxicology*
Manuscript ID	JAT-16-0251.R2
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Callegaro, Giulia; Universita degli Studi di Milano-Bicocca Dipartimento di Scienze dell\'Ambiente e del Territorio e di Scienze della Terra Corvi, Raffaella; European Commission Joint Research Centre Ispra Sector, European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM), Systems Toxicology Unit Salovaara, Susan; European Commission Joint Research Centre Ispra Sector, European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM), Systems Toxicology Unit Urani, Chiara; Universita degli Studi di Milano-Bicocca Dipartimento di Scienze dell\'Ambiente e del Territorio e di Scienze della Terra, Stefanini, Federico; Università di Firenze, Statistica, Informatica, Applicazioni
Keywords:	Cell Transformation Assay, BALB/c 3T3 cell line, Nickel chloride, Statistical image descriptors, Dose-response, Foci, Benzo[a]Pyrene, Extended linear mixed-effects models

Relationship between increasing concentrations of two carcinogens and statistical image descriptors of *foci* morphology in the Cell Transformation Assay

Giulia Callegaro[1], Raffaella Corvi[2], Susan Salovaara[2], Chiara Urani[1], Federico M. Stefanini[3]

[1] Department of Earth and Environmental Sciences, University of Milan Bicocca – Piazza della Scienza, 1 – 20126 Milan, Italy

[2] European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM), Chemical Safety and Alternative Methods Unit; Directorate F, European Commission Joint Research Centre, TP 126, Via E. Fermi 2749, I-21027 Ispra (VA), Italy

[3] Department of Statistics, Computer Science, Applications, University of Florence - Viale Morgagni 59, 50100 Florence, Italy

Running head: Effects of carcinogens concentration on *foci* morphology

Chiara Urani and Federico M. Stefanini are co-last Authors

Corresponding Author:
Federico Mattia Stefanini
Department of Statistics, Computer Science, Applications
University of Florence
Viale Morgagni 59, 50100 Florence, Italy
Phone: +39 0554237266
stefanini@disia.unifi.it

Abstract

Cell transformation assays (CTAs) have long been proposed for the identification of chemical carcinogenicity potential. The endpoint of these *in vitro* assays is represented by the phenotypic alterations in cultured cells which are characterized by the change from the non-transformed to the transformed phenotype. Despite the wide fields of application and the numerous advantages of CTAs, their use in regulatory toxicology has been limited in part due to concerns about the subjective nature of visual scoring, i.e. the step in which transformed colonies or *foci* are evaluated through morphological features.

An objective evaluation of morphological features has been previously obtained through automated digital processing of *foci* images to extract the value of three statistical image descriptors. In this study a further potential of the CTA using BALB/c 3T3 cells is addressed by analysing the effect of increasing concentrations of two known carcinogens with different mode of action (Benzo[a]Pyrene and $NiCl_2$) on *foci* morphology.

The main result of our quantitative evaluation shows that the concentration of the considered carcinogens has an effect on *foci* morphology that is statistically significant for the mean of two among the three selected descriptors. Statistical significance also corresponds to visual relevance.

The statistical analysis of variations in *foci* morphology due to concentration allowed to quantify morphological changes that can be visually appreciated but not precisely determined. Therefore, it has the potential of providing new quantitative parameters in CTAs, and of exploiting all the information encoded in *foci*.

Short abstract for Table of Contents

The cell transformation assays (CTAs) have long been proposed for the identification of chemical carcinogenicity potential. The endpoint of CTAs is the induction of *foci* of transformed cells, characterized by specific morphological features, after treatment with a suspected carcinogen.

In this study *foci* morphology is shown to be statistically related to the concentration of two known carcinogens (Benzo[a]Pyrene and $NiCl_2$), as captured by selected statistical image descriptors.

Statistical descriptors may represent new quantitative parameters for the assessment of cell transformation.

Keywords:

- Cell Transformation Assay
- BALB/c 3T3 cell line
- *Foci*
- Nickel chloride

- Benzo[a]Pyrene
- Statistical image descriptors
- Dose-response
- Extended linear mixed-effects models

Abbreviations:

BD, Boundary Index

BIC, Bayesian Information Criteria

B[a]P, Benzo[a]Pyrene

C, concentration

CTA, *in vitro* Cell Transformation Assay

ED, Equivalent Diameter

EFP, perimeter of the circle with equivalent area

FP, *focus* perimeter

MD, Median

PD, Petri dish

1. INTRODUCTION

The cell transformation assays (CTAs) have long been proposed for the identification of chemical carcinogenicity potential. The endpoint of these *in vitro* assays is the induction of morphological alterations in cultured cells, which are characterized by the change from the non-transformed to the transformed phenotype. Long ago Barrett and Ts'o (1978) demonstrated that the study of oncogenesis in cell culture as a model for neoplastic transformation *in vivo* is strongly justified, and that the induction of specific morphological alterations in cultured cells reflects the features of cells exhibiting neoplastic potential *in vivo*.

These assays offer several advantages in comparison to the *in vivo* bioassays in rodents (OECD, TG451). They *i*) are faster and more cost efficient; *ii*) allow the identification of not only genotoxic, but also of some non-genotoxic compounds; *iii*) may help to clarify *in vitro* genotoxic positive results; *iv*) provide a means to investigate tumor promotion activities and efficacy of chemopreventive agents; and *v)* support the 3Rs principles of replacement, reduction and refinement (EURL ECVAM, 2012). Current users of CTAs are the industry (chemical, agro-chemical, pharmaceutical, and tobacco) for screening purposes, and academia mainly to investigate the molecular mechanisms of carcinogenesis. Although the use of CTAs is not currently an explicit regulatory requirement in any area of toxicology, their application is mentioned in different contexts in a number of guidance documents (EURL ECVAM Recommendation, 2013). In addition, two guidance documents on the CTA performed using Syrian hamster embryonic (SHE) cells and the Bhas 42 cell line, respectively, have been adopted by the Organization for Economic Cooperation and Development (OECD), which will probably stimulate the broader use of the assay in the future (OECD, 2015; 2016).

Despite the wide fields of applications and the numerous advantages of CTAs, their use in regulatory toxicology has been limited in part due to concerns about the reproducibility of the results between different laboratories in relation to the subjective nature of using morphological features for the identification of transformed cells, and the lack of understanding the mechanisms underlying the process of transformation (EURL ECVAM, 2012).

Recently several efforts have been undertaken to standardize and validate this test method, as well as to improve and refine different aspects of these assays. In particular, work is underway to respond to concerns that have been raised, such as the need for an objective scoring of transformed *foci*. This was addressed in the prevalidation and validation studies coordinated by the European Reference Laboratory on Alternative Methods (EURL ECVAM) and the Japanese Centre for the Validation of Alternative Methods (JaCVAM), which led to the publication of standardized protocols in combination with photo catalogues to support the researchers during the scoring

procedure (Bohnenberger et al., 2012, Maire et al, 2012; Sasaki et al., 2012a; Sasaki et al., 2015; OECD, 2016 Annex 2). In the framework of this study, a new statistical method to improve the analysis of BALB/c 3T3 CTA data was developed (Hoffmann, 2012). Moreover, efforts addressed at the automation of the visual scoring and of *foci* classification by means of image analysis are ongoing in our group and several approaches based on extraction of statistical descriptors have already been published (Ridder et al., 1997; Procaccianti et al., 2011; Urani et al., 2009, 2013; Callegaro et al., 2015). In these studies, digital images of *foci* were acquired and regions of interest corresponding to every single *focus* were identified. Statistical descriptors reflecting the coded morphological features, which are at the basis of visual scoring of *foci*, were developed with the aim of building a quantitative *foci* classifier.

The goal of this study is to provide the means to further exploit CTA results in order to consider the information encoded in *foci* morphology. To this aim, the effect of different concentrations of two known carcinogens, Benzo[a]Pyrene (B[a]P) and nickel chloride ($NiCl_2$) on *foci* morphology is assessed in the CTA using BALB/c 3T3, as captured by three selected *foci* descriptors. As far as we are aware, this is the first attempt to quantitatively evaluate the role played by different concentrations of carcinogens on the morphological endpoint in the transformation process. The two carcinogens were selected based on their different modes of action, (B[a]P) which exerts its effects through genotoxic mechanisms, and $NiCl_2$ as it is known to act through indirect interactions with DNA.

2. MATERIALS AND METHODS

2.1. CTA experiments

The BALB/c 3T3 CTA experiments were previously performed by EURL ECVAM within two studies. The first study aimed at evaluating the transforming properties of different forms of nickel (Salovaara et al., in preparation) and the other one assessed the transforming potential of B[a]P as part of a broader study aimed at prevalidating the BALB/c 3T3 CTA (Corvi et al., 2012; Sasaki et al., 2012b). The assays were carried out according to the experimental protocol used in the EURL ECVAM prevalidation study (Corvi et al., 2012; Sasaki et al., 2012b). A concentration of 4 µg/ml of 3-Methylcholanthrene (MCA, CAS number #56-49-5) was chosen as the positive control as this concentration was reported to induce positive transformation responses (Tanaka et al., 2012). In the present work we focused on the dishes treated with different concentrations of $NiCl_2$ (CAS number #7718-54-9**),** more precisely the tested concentrations were 250, 275, 300, 350 and 400 µM. (CAS number #50-32-8) Tested concentrations for B[a]P (CAS number #50-32-8) , were 0.0005, 0.005, 0.001, 0.05, 0.125, 0.625, 3.125, 15 µg/ml.

As described in the recommended protocol and in the photo catalogue, the scoring of transformed *foci* was performed under a stereomicroscope according to predefined criteria related to morphological characteristics, which allowed to assign the *foci* into three categories: Type I, Type II and Type III (Tanaka et al. 2012; Sasaki et al., 2012a). Only Type III *foci* were considered as transformed, and thus scored as positive. According to the recommended protocol, a *focus* with a diameter of less than 2 mm is too small to be considered as transformed.

Type I *foci* are small, non-invasive and weakly basophilic. Type II *foci* are multilayered, contain basophilic spindle shape cells, and exhibit some cell piling up and criss-crossing, although to a limited extent. Type III *foci* show dense multilayering (piling up) and display deep basophilic spindle-shape cells, which differ from the non-transformed monolayer of background cells, that shows contact-inhibition, and upon confluence is composed of flat, epithelial-like cells.

2.2 Workflow

The effect of different concentrations of two carcinogens on *foci* morphology was investigated through image analysis, as described by the workflow in Figure 1.

After performing the experiments (Box 1), *foci* images were acquired under a stereomicroscope equipped with a digital camera (Box 2), and statistical image descriptors were calculated for each *focus* (Box 3). A selection of three previously developed descriptors (Urani et al., 2013) was used to summarize the morphological features that are recognized as relevant in visual scoring (Sasaki et al., 2012a). As a result, a database comprising descriptors of *foci* morphology was built (Box 4), in which each collected *focus* is stored together with the value taken by its three descriptors, the carcinogen concentration and further auxiliary information represented by the label of the Petri dish in which the *focus* was acquired. The methodologies used in these steps, shown in Figure 1 by Boxes 2, 3, and 4, are detailed by Callegaro et al. (2015) and summarized in *section 2.3*. The effect of concentration on *foci* morphology, as captured by selected image descriptors, was studied by developing extended linear mixed-effects models (Pinheiro and Bates, 2000, p.202 eq. 5.1) in which the correlation existing among *foci* collected from the same Petri dish was also taken into consideration. In each final model (3 descriptors times 2 carcinogens = 6 final models), a statistical test of the hypothesis stating the null effect of carcinogen concentration on *foci* morphology was performed. Further details about the statistical analysis in Boxes 5b and 5c of Figure 1 are provided in *section 2.4*.

2.3. Foci digitizing and data acquisition

Foci images were acquired by adopting the procedure previously developed by Callegaro et al. (2015), and briefly described below.

Acquisition was performed under a stereomicroscope (Zeiss, Stemi SV6) equipped with 6.3× lens (Carl Zeiss, Arese, Italy) and a digital camera (AxioCam Mrc5, 36 bit). Images were saved in TIFF-48bit RGB format and had a size of 2572 × 1928 pixels, where 1 pixel is equivalent to a real size of 6.7842×10^{-3} mm (1 cm = 1474 pixels). Only fully transformed Type III *foci* were considered for the acquisition.

Foci images comprised both the region corresponding to the transformed *focus* and a fraction of the surrounding monolayer of non-transformed cells. An original segmentation algorithm coded in Python (Python Software Foundation) was developed by Callegaro et al. (2015) in order to isolate the *focus* regions from the surrounding monolayer. The algorithm acquires a RGB colour space *focus* image, converts it into the HSV colour space and separates the region corresponding to each *focus* from the surrounding monolayer by setting appropriate intensity thresholds.

During the segmentation process a subset of all images (around 30% for $NiCl_2$ dataset, around 40% for B[a]P dataset) was discarded, mainly due to pen marks used in the visual scoring process. The derived $NiCl_2$ dataset comprised a collection of 165 Type III *foci* images, obtained from dishes tested with all the five $NiCl_2$ concentrations specified in *section 2.1*. While the B[a]P derived dataset comprised 224 Type III *foci* images, obtained from dishes tested with all eight B[a]P concentrations (see *section 2.1*).

The experiment has a balanced layout as regards the number of Petri dishes: ten Petri dishes were available for each treatment (i.e. $NiCl_2$ or B[a]P concentration). Nevertheless, a different number of Type III *foci* was observed in Petri dishes treated at different concentrations of the same carcinogen, therefore the resulting sample size is not constant over concentrations. The change in the number of segmented *foci* at each concentration was also due to the pen marks on Petri dishes, but no appreciable changes in the trend of number of *foci* depending on concentration are produced, as can be seen in the graphs of Figure 2 and Figure S1. See Table 1, Table 2 and Figure 2 for details on the structure of the two datasets after segmentation. Additional information about the number of *foci* observed at different concentrations before segmentation is shown in Figure S1.

For each image region corresponding to a Type III *focus* included in the final database after segmentation, three previously developed statistical descriptors were calculated (Urani et al., 2013). These descriptors summarize three morphological features among those coded and visually scored in BALB/c 3T3 CTA (Sasaki et al., 2012b), which are:

- Equivalent Diameter (ED);
- Median (MD);
- Boundary Index (BD).

The descriptor ED is defined as the diameter of the circle whose area is equal to the area of the *focus*: $ED=2\left(\sqrt{(area/\pi)}\right)$, where π is the trigonometric constant. We selected the ED descriptor to

evaluate the size of the *focus*, taking into account the foci's shape polymorphism.

The descriptor MD is the median of the distribution made by image grey-levels in the region corresponding to a *focus.* The median value of this distribution depends on the degree of multilayer growth. Given that the intensity of light source is constant, image grey-levels mostly depend on the amount of light passing through the *focus* region: regions with dense piling up will correspond to darker image pixels, while regions that are less multilayered will correspond to lighter pixels.

Finally, the BD descriptor takes high values when invasive growth in the surrounding monolayer is pronounced, a typical feature of fully transformed cells belonging to Type III *foci.* Invasiveness often implies boundary heterogeneity of *foci*, thus the BD index compares the actual perimeter of the *focus* (FP) with the perimeter of the circle of equivalent area (EFP): $BD = \left(\frac{FP}{EFP}\right) - 1$.

The above three descriptors were previously developed and applied to the analysis of *foci* images obtained from the prevalidation study coordinated by EURL ECVAM (Tanaka et al., 2012; Urani et al, 2013), which included a large amount of experiments conducted with coded and uncoded chemicals.

2.4. Statistical analysis

A generic Type III *focus* in a given Petri dish x_{PD} is characterized by the random variables Y_{ED}, Y_{MD}, Y_{BD} describing the value taken by statistical descriptors ED, MD and BD. The joint distribution of Y_{ED}, Y_{MD}, Y_{BD} given the Petri dish x_{PD} and the concentration x_C was considered at the start of model building. The factorization of the joint distribution function into univariate conditional distributions also allowed to study the dependence existing among statistical descriptors:

$$p(y_{ED}, y_{MD}, y_{BD} | x_{PD}, x_C, \theta) = p(y_{ED}|, y_{MD}, y_{BD}, x_{PD}, x_C, \theta_{ED}) \cdot p(y_{MD}|y_{BD}, x_{PD}, x_C, \theta_{MD}) \cdot p(y_{BD}|x_{PD}, x_C, \theta_{BD}) \quad \text{eq. 1}$$

where $\theta = (\theta_{ED}, \theta_{MD}, \theta_{BD})^T$ is the partitioned vector of model parameters. In Figure 3, the initial factorization for one single *focus* is shown by a Directed Acyclic Graph (details in Buntine, 1994) in which model parameters are omitted. Each conditional distribution on the right of equation (1) was parameterized assuming a Normal response explained by conditioning variables. Thus the starting model was in the class of linear mixed-effects models (Pinheiro and Bates, 2000) for each statistical descriptor. The concentration values were rescaled in the range [1,3] for the B[a]P carcinogen. The class of investigated extended linear mixed-effects models included linear predictors made by low degree polynomials of explanatory variables and their interactions, while the variance of the response was modelled as a function of the (rescaled) concentration. Model fitting was performed by Maximum Likelihood, after considering *foci* as conditionally exchangeable. A preliminary selection of reasonable models was performed looking for models

with a small value of the Bayesian Information Criterion (BIC), while a Likelihood ratio test (Cox and Hinkley, 1974) for the effect of concentration was performed by comparing models with/without the explanatory variable concentration.

The final model of each descriptor belongs to the class of extended linear models (Pinheiro and Bates, 2000, p.202 eq. 5.1). The within group correlation existing among values of a descriptor assessed on *foci* located in the same Petri dish was computed by exploiting an R object in the class corCompSymm (nlme package, Pinheiro et al., 2015): we considered the Petri dishes as separate environments, characterised by micro-fluctuations of experimental conditions shared by all *foci* in each dish.

Residuals calculated from the final model of each image descriptor were graphically inspected to check model assumptions. Details about model diagnostic are provided as Supporting Information, e.g. quantile-quantile plots of residuals for each model (Figures S2-S3).

All statistical computations were performed in R (R Core Team, 2012), using *nlme* (Pinheiro et al., 2015) and *effects* packages (Fox, 2003).

2.4.1. The factorization of the joint distribution

The statistical image descriptors ED, MD and BD carry information about morphological features jointly used for *foci* scoring, as previously shown by Urani et al. (2013), therefore the correlation among them was expected. The configuration of considered dependencies is represented in Figure 3. The blue directed arrows shown in the graph represent dependence between the considered descriptors (details in Buntine, 1994). The influence of carcinogen concentration on the selected statistical descriptors is represented by yellow directed arrows starting in node C (Figure 3), while yellow arrows leaving node PD and reaching BD, ED and MD indicate the effect of Petri dish environment on *foci* morphology.

The factorization shown in equation 1 (*section 2.4.*) was selected by taking into account the specific features of each descriptor before looking at the data. A *focus* can expand in the x-y space of the Petri dish, determining its size, but it also grows in the third direction, the z axis of the Petri dish. Hence, it is expected that the MD descriptor would be also informative for the ED descriptor, which describes the dimension of a *focus*. Furthermore, the ability to invade the surrounding monolayer of non-transformed cells can be considered as influencing the growth process in the two-dimensional space of the Petri dish. Accordingly, the dependence relationship between ED and BD descriptors has to be taken into account (blue arrows in Figure 3).

Transformed BALB/c 3T3 cells in a *focus* exhibit different growth characteristics from those of non-transformed cells of the surrounding monolayer, primarily represented by a loss of contact-inhibition and uncontrolled proliferation due to complex molecular mechanisms. In a *focus*

displaying these properties, the invasion on the monolayer of non-transformed cells can be observed, often by means of prolonged arms and vortexes of polarized cells (see for example figure 7D of Urani et al., 2013). This behaviour of uncontrolled growth and invasiveness influences the *in vitro* three-dimensional growth (multilayering) of *foci*. Thus we developed a model for the response of the MD descriptor in which BD is an explanatory variable.

In addition, each Petri dish represents a separate environment in which *foci* are exposed to the same micro-environmental fluctuations and experimental errors, in contrast to *foci* observed in different dishes. For this reason, we have included into the model an intra-class correlation parameter (Pinheiro and Bates, 2000) which describes the degree of resemblance of *foci* in the same Petri dish for a given descriptor, as detailed in *section 2.4*. Note that, in the final model, we did not estimate the effects of Petri dishes since not relevant for the estimate of the effect of concentration on *foci* morphology, thus reducing the number of model unknowns.

2.4.2. Models

The estimated expected values of the three conditional distributions for the $NiCl_2$ and B[a]P carcinogens are shown respectively in Table 3 and 4, just before testing for the contribution of concentration, therefore the concentration is included as explanatory variable.

3. RESULTS

3.1. Conditional models of statistical descriptors given the concentration of a carcinogen

Considering the factorization of the *equation 1* and represented in Figure 3, in this section the final models of each image descriptor given a specific carcinogen are described and commented. For details about model fitting, see *sections 2.4.1* and *2.4.2*.

3.1.1 The Equivalent Diameter does not depend on carcinogen concentration

Models built to test the dependence of the expected value of ED on carcinogen concentration are indicated as model N1 ($NiCl_2$ dataset) and model B1 (B[a]P dataset) and they both include the other two descriptors, MD and BD, as explanatory variables. The model equations are shown in Table 3 and 4, first row. A remarkable feature of both models is that the effect of concentration on ED after including MD and BD was not found statistically significant. For model N1, the likelihood ratio test did not lead to the rejection of the hypotheses stating the null effect of concentration, with a p-value equal to 0.2427. Likewise, for model B1, the likelihood ratio test did not lead to the rejection of the hypothesis stating the null effect of concentration, with a p-value equal to 0.8781. It is worth noticing that, although the expected value of ED does not depend on concentration, the value of variance does. Therefore, the conditional distribution of ED of both models indeed depends on concentration.

In practical terms, the average size of a *focus* does not depend on the concentration of a carcinogen when the comparison is performed among *foci* with the same values of BD and of MD.

Graphical summaries from the fitted models are shown in Figure 4, panels C and H, where linear relationships between the *focus* size and the concentration of the treatment can be observed. While the expected values for model B1 vary with B[a]P concentration according to a straight line whose slope is close to zero (panel H), for model N1 ($NiCl_2$ carcinogen) the correspondent slope is negative instead (panel C). Still, testing for the partial contribution exerted by concentration as explanatory variable led to the conclusion that concentration is not significant, as suggested also by the overlapping confidence intervals in Figure 4H. For each model, four examples of *foci* are provided in the same figure, whose values of ED can be found in the extreme regions of the graph: for model N1 in Figure 4, panels A, B, D and E, for model B1 in the same Figure, panels F, G, I, J. Regardless the carcinogen concentration, small *foci* having low values for the ED descriptor (Figures 4, panels A and D, F and G) or bigger *foci* showing higher values for the ED descriptor (Figure 4, panels B and E, I and J) can be observed.

The dependence relationship between ED and MD descriptors was further explored by elaboration from fitted models. Considering the models presented here (see Table 3 and 4), a general large *focus*

is characterized by a substantial ability to grow into multilayers, as shown by the *focus* in Figure 4, panels B and E, G and J. A reduced multilayer growth, in contrast, can be found in small *foci*, as the four provided in Figure 4, panels A and D, F and I.

Keeping in mind that MD is directly proportional to the amount of light in the *focus* grey-level image, low values for the MD descriptor will correspond to darker *foci*, thus *foci* characterised by several layers of cells. While high values for the MD descriptor will be associated with less multilayered *foci.*

Indeed, an inverse relationship can be found plotting the fitted values of ED *versus* MD descriptor, as provided in Supporting Information, Figure S2 and S3, panels G.

A remarkable result from our analyses is that we found the same qualitative relationships among descriptors and of the concentration on descriptors, despite that models were fitted to data from two different types of carcinogen, one genotoxic and one non-genotoxic. Thus it will be interesting in the future to perform similar studies on additional carcinogens to test the general validity of such qualitative features.

3.1.2 The Median is dependent on carcinogen concentration

Models built to test the dependence of the expected value of Median descriptor (MD) on carcinogen concentration are indicated as Model N2 ($NiCl_2$ dataset) and Model B2 (B[a]P dataset) and they both include BD as explanatory variable (see Figure 3). The model equations are shown in Table 3 and 4, second rows.

In both models, the partial contribution of concentration after including BD into the model is statistically significant: for model N2, the likelihood ratio test for the hypothesis of no effect of concentration was rejected with a p-value lower than 0.005 (p-value 0.0025). Likewise, for model B2, the likelihood ratio test for the hypothesis of no effect of concentration was rejected with a p-value lower than 0.001 (p-value 0.0007). Hence for the average *focus*, multilayered growth changes with the carcinogen concentration when the comparison is performed among *foci* characterized by the same value of BD.

With $NiCl_2$, the best model (N2) includes a second-degree polynomial in the concentration, thus the change of MD's average is not linear with respect to the increase of concentration (Figure 5, panel B).

Figure 5B describes the U-shaped non-linear behaviour of the MD descriptor obtained in Petri dishes treated with increasing concentrations of $NiCl_2$. A U-shaped trend can be observed with a minimum located at about 310 μM.

As MD is directly proportional to the amount of light in the *focus* grey-level image, high values for the MD descriptor will thus correspond to lighter *foci*, characterised by a reduced multilayering

(Figure 5C). It follows that low values for the MD descriptor will be associated to *foci* characterized by a significant multilayer growth, as the one presented in Figure 5D.

In the bounded range of considered concentrations, a first relative maximum, corresponding to 250 μM $NiCl_2$, is related to medium MD values, corresponding to *foci* not markedly and/or uniformly multilayered (see as example *focus* in Figure 5A); moderately higher MD values can be observed corresponding to the highest concentration (400 μM $NiCl_2$), as illustrated by the *focus* in Figure 5C. In between, corresponding to intermediate $NiCl_2$ concentrations (275-350 μM), the absolute minimum of the curve plotted in Figure 5B is located, suggesting the presence of multilayered *foci* (as e.g. Figure 5D).

The best model for MD is not monotonic with the increase of concentration: higher and lower concentrations induce the formation of *foci* displaying a reduced multilayer growth, while *foci* obtained from intermediate $NiCl_2$ concentrations show an increase in this feature.

By contrast, in the best model built for B[a]P, the linear predictor is a straight line with respect to the concentration (Figure 5, panel F). Considering again that MD is directly proportional to the amount of light in the *focus* grey-level image, at low B[a]P concentrations *foci* having small MD values are found (Figure 5E), in contrast to the higher MD values observed in *foci* at the highest concentration.

Even if the two models differ in the shape of the curve representing the expected values in relation to the concentrations, they partially share a similar feature, which is that the expected value of MD is high at very high concentrations of carcinogens.

We conjecture that the trend at very high concentrations, shared by both carcinogens, could be due to a combination of three factors. Firstly, these very high concentrations might have cytotoxic effects, dampening the multilayer growth of the transformed *foci* in these Petri dishes. This conclusion is supported by the fact that these Petri dishes had fewer *foci*, when compared to lower doses (Tables 1 and 2, Figure 2). Secondly, in *foci* induced by the highest concentrations used, a remarkable heterogeneity, a hallmark of both intermediate and mixed *foci*, was found as shown in Figure 5G. The MD descriptor applied to such a variable region, could have underestimated the overall multilayer growth of *foci* found in Petri dishes treated with high doses. From the statistical point of view, the presence of cytotoxic effects of carcinogens possibly responsible for the decreased number of *foci* observed at high concentrations, could in turn have caused a reduced precision (and statistical power) of parameter estimates.

The expected value of MD with respect to the concentration differs in the two models, in particular in the low range of concentration, which could be due to specific differences of the carcinogens mode of action. Nonetheless, the concentration range of B[a]P was experimentally designed to be equally spaced on the logarithmic scale, hence a quite large part of the inner range has no

observations on the original scale by design. Thus, there is a certain degree of uncertainty related to the estimate at intermediate range of concentrations, therefore the apparent linear shape could not be confirmed in specifically designed experiments with more distinct concentrations. Still, the concentration exerts a statistically significant effect on the MD descriptor.

In Figures S2 and S3, further elaborations based on the fitted models (Model N2 and Model B2) show the dependence relationship of MD from BD, in particular a monotonic relationship was found in both cases. From these models, we expect that highly multilayered *foci* will be less invasive than less multilayered *foci*.

3.1.3 The Boundary Index is dependent on carcinogen concentration

Models built to test the dependence of Boundary Index (BD) on carcinogen concentration are indicated as Model N3 and Model B3, respectively for $NiCl_2$ and B[a]P datasets. The effect of concentration on BD was statistically significant in both models, and the expected value of BD is linear in the concentration (see Figures 6B and 6E): high values of BD result from high concentrations. In particular, for model N3, the likelihood ratio test for the hypothesis of no effect of concentration was rejected with a p-value lower than 0.001 (p-value 0.0004), while for model B3, the p-value of the test was lower than 0.0001.

It is noteworthy to remember that the BD index is related to the degree of departure of the *focus* shape from that of a perfect circle with an equivalent area. Finger-like protuberances can be seen at the edge of the *focus*, as a result of the criss-cross growth at its edges (Sasaki et al., 2012a). The BD descriptor captures a trait of invasive growth that is characterized by heterogeneous growth at a *focus* boundary (Urani et al., 2013), therefore at higher concentrations of carcinogen more pronounced finger-like protuberances of *foci* are expected.

After treatment with the lower concentrations of both carcinogens (250-300 µM of $NiCl_2$ and from 0.0005 to 3.125 µg/ml of B[a]P), as described in *section 3.1.2*, fairly multilayered *foci* can be frequently observed: this phenotype is often associated with a circle-like growth (see as examples Figures 6A and 6D).

At higher concentrations, 350-400 µM $NiCl_2$ and 15 µg/ml of B[a]P, *foci* showing a more fragmented morphology can be found, such as those presented in Figures 6C and 6F. Within the same *foci,* distinct regions of aggregation can be observed, characterized by tightly packed spindle-shaped cells.

We remark, once more, that despite the different nature of the two carcinogens, qualitatively similar models for BD as a function of concentration were obtained.

4. DISCUSSION

In a previous study (Urani et al., 2013) three statistical descriptors of *foci* morphology were developed with the aim of mimicking those features that are assessed in visual scoring in CTAs for chemical carcinogenicity testing. In this work, we selected $NiCl_2$ and B[a]P *foci* images from previously performed CTAs, and we developed statistical models to test the effect of concentration on morphology. The joint distribution of all three statistical descriptors (ED, MD and BD) was factored into the product of three conditional distributions, after recognizing the statistical dependences existing among descriptors. Furthermore, we also considered that *foci* in the same Petri dish could be much more similar than *foci* located in different Petri dishes, due to the shared micro-environment.

This study suggests that the statistically significant effects of carcinogen concentration on some features of *foci* morphology, as represented by two of the three selected descriptors, are also coupled to visual relevance. *Foci* images obtained by testing two carcinogens with different mode of action were analyzed with the same method, and similar results in the structure of the dependencies with the concentration were obtained. Indeed, by inspecting the morphology of the *foci* obtained for each carcinogen at different concentrations, an appreciable degree of visual dissimilarity has been observed. At lower concentrations of both carcinogens (250-300 µM $NiCl_2$ and 0.0005-3.125 µg/ml B[a]P) analyzed *foci* seem to grow on average in multilayers and in a circle-like shape. While at higher concentrations (300-400 µM $NiCl_2$ and 15 µg/ml B[a]P) the *foci* show a fragmented morphology characterized by less uniform multilayer growth. It must be pointed out that the ranges of $NiCl_2$ and B[a]P were preliminarily selected by dose-range finding tests, as dictated by the CTA standard protocol, and qualitative similar shapes of boxplots for the number of *foci* in the concentration ranges were observed. By contrast, no clear differences can be seen in *foci* dimension in relation to the concentrations when the comparison is performed among *foci* with the same values of BD and of MD. In fact, as demonstrated by our statistical analyses, the concentration does not exert a direct effect on ED when MD and BD are already taken into account. In other terms, the expected value of ED does not change with concentration if *foci* of similar MD and BD are compared. It must be remarked that the carcinogen concentration has a significant effect on some aspects of *foci* morphology even if the concentration is included into the model as a qualitative factor (results not shown), that is without imposing a specific shape on the concentration-to-descriptor relationship. This is true for both the carcinogens considered. Note also that the variance of ED depends on the concentration, thus the variability changes with concentration.

In the models developed here, the dependence among descriptors and of a descriptor on concentration is not always linear. This may reflect the inherent biological complexity of the

carcinogenic process. In particular, such non-linearity was evidenced in the $NiCl_2$ model. Nickel compounds and B[a]P are classified as Group I carcinogens (IARC, 2012a; IARC, 2012b) affecting human health through occupational and environmental exposure. While B[a]P is known to act through a genotoxic mechanism that involves biotransformation to highly reactive metabolites that form covalent adducts to DNA and other genotoxic effects, carcinogenicity induced by nickel compounds is characterized by the induction of oxidative stress through generation of reactive oxygen species and by the interference with DNA repair pathways, thus leading to genetic instability. Furthermore, nickel interferes with DNA methylation and histones acetylation, and activates hypoxic signaling. These mechanisms taken together cause deregulation of cell proliferation (Cameron et al., 2011; Sun et al., 2013; Magaye et al., 2012). As suggested by Haber et al. (2000), the indirect interaction with DNA in soluble nickel-induced *in vitro* transformation could imply a non-linear dose-response relationship. However, the authors also remarked that the overall available data are insufficient to determine the doses at which such non-linearity occurs. Nonetheless, they also underline that the suggestion of such non-linearity is consistent with the negative animal carcinogenicity studies for soluble nickel, despite this negative results could also be due to a different bioavailability and clearance of nickel ions after exposure to soluble nickel compounds, as recently published (Goodman et al., 2011). As far as we are aware, no major developments have occurred in this direction. In our dataset the effect of $NiCl_2$ on one image descriptor (MD) was here estimated to be non-linear as a function of concentration. Widening the study of statistical descriptors to a broader concentration range might support the non-linearity hypothesized by Haber et al. (2000), and in such expanded range they could be studied in a quantitative way, thus possibly defining the concentration range at which non-linearity occurs.

In vitro cell transformation is considered to be a result of stepwise genotypic alterations, which underlie the corresponding phenotypical ones (Smets, 1980; Keshava, 2000). Increasing concentrations of carcinogens may lead to an accumulation of genetic changes (or different steps of the global process of transformation), resulting in a variety of phenotypes which reflect different molecular alterations. Quantitative morphological descriptors discriminate between the variety of morphologies that can be visually appreciated but not precisely determined, leading to a detailed evaluation of transformed *foci.* This may be especially useful to reduce the uncertainty in the classification of mixed or intermediate *foci*.

Our analysis showed for the first time the dependence of Type III *foci* morphology on concentration, thus it enabled the possibility of further distinctions inside the usual classification scheme, as previously suggested by other authors (Keshava, 2000; Lu et al., 1986). The utility of these detectable differences needs further investigations. Specific experiments could be designed to evaluate a potential correspondence between different morphologies in Type III class *foci* and

different tumor formation probabilities. Similarly, the molecular basis of observed differences within Type III *foci* should be further characterized to determine, for example, whether differences due to concentration detected by image descriptors are associated to specific molecular/metabolic signatures. Ongoing research points towards the molecular comprehension of *in vitro* cell transformation (Forcella et al., 2016; Priya et al., 2013; Ao et al., 2010; Rohrbeck et al., 2010; Walsh et al., 2009; Clemens et al., 2005). The potential availability of quantitative morphological markers representative in transformed cells of specific molecular pathways and specific classes of damages as a function of concentration could open the way towards an improvement of CTA.

Further applicative relevance of our approach includes the possibility to predict the values of the three descriptors summarizing the morphological features of *foci* in the CTA for other concentrations comprised in the tested range but not actually experimentally assessed.

Additional improvements of the proposed method can be envisaged. The fraction of all *foci* images lost in the segmentation step could be reduced in part by specific experimental precautions, for instance by limiting the use of pen marks outside the area of *foci*. In addition, improvements of the proposed descriptors might extract further quantitative information which is possibly relevant for the assessment of transformed *foci*. For example, the proposed Boundary Index (BD) captures just one of the aspects of invasive growth, in particular the degree of departure of the *focus* shape from that of an exact circle. Low values for the BD index seem to be related to *foci* composed of separated regions of cellular aggregates, rather than a unique *focus* body. A descriptor taking into account the heterogeneity within the same *focus* could improve both the assessment of invasiveness and the characterization of *foci* with mixed or intermediate morphology. Finally, large screening tests are needed for a detailed characterization of the effect of concentration on *foci* morphology, given that we provided positive evidence for just one compound for each major class of carcinogens (genotoxic and non-genotoxic).

The method presented in this study and applied to the BALB/c 3T3 CTA, could be potentially useful also in the CTA with Bhas 42 cells, for which an OECD Guidance Document has recently been published to support its use in regulatory applications (OECD, 2016, Sakai et al., 2011). The Bhas 42 cell line is in fact derived from BALB/c 3T3 cells transfected with the *v-Ha-ras* proto-oncogene, consequently its transformed *foci* are characterized by the same morphological features as the BALB/c 3T3 *foci*.

In conclusion, our quantitative evaluations mainly show that the concentration of the considered carcinogens exerts an effect on *foci* morphology. We believe that statistical descriptors of *foci* morphology have the potential of mimicking trained human scorers very well, but with the advantage of being both objective and quantitative.

In this way, additional information can be extracted from CTAs to be exploited in the application of

more reliable *in vitro* procedures when assessing carcinogenicity through an integrated testing strategy, either in a full probabilistic (Stefanini, 2013) or a Weight of Evidence approach (ECHA 2010).

Acknowledgments

This work is partially funded by the University of Milan Bicocca (2014-ATE-0043 to CU) and by the University of Florence (to FMS), funding framework "Strategic project of basic research, year 2015." The authors acknowledge the European Commission.

REFERENCES

Ao L, Liu JY, Liu WB, Gao LH, Hu R, Fang ZJ, Zhen ZX, Huang MH, Yang MS, Cao J. 2010. Comparison of gene expression profiles in BALB/c 3T3 transformed *foci* exposed to tumor promoting agents. *Toxicol. In Vitro.* **24**: 430-438.

Barrett JC and Ts'o PO. 1978. Evidence for the progressive nature of neoplastic transformation in vitro. *PNAS.* **75** (8): 3761-3765.

Bohnenberger S, Kunkelmann T, Perschbcher S, Poth A, Pant K, Bruce SW, Sly JE, Schwind KR. 2012. Photo catalogue for the classification of cell colonies in the Syrian hamster embryo (SHE) Cell Transformation Assay at pH 6.7. *Mutat. Res.* **744** (1): 82-96.

Buntine WL. 1994. Operations for Learning with Graphical Models. *J. Artif. Intell. Res.* **22**: 159-225.

Callegaro G, Stefanini FM, Colacci A, Vaccari M, Urani C. 2015. An improved classification of *foci* for carcinogenicity testing by statistical descriptors. *Toxicol. In Vitro*. **29** (7): 1839-50.

Cameron KS, Buchner V and Tchounwou PB. 2011. Exploring the Molecular Mechanisms of Nickel-Induced Genotoxicity and Carcinogenicity: A Literature Review. *Rev. Environ. Health.* 2011, **26**(2): 81–92.

Clemens F, Verma R, Ramnath J, and Landolph JR. 2005. Amplification of the Ect2 proto-oncogene and over-expression of Ect2 mRNA and protein in nickel compound and methylcholanthrene-transformed 10T1/2 mouse fibroblast cell lines. *Toxicol. Appl. Pharmac.* **206**, 138–149.

Corvi R, Aardema MJ, Gribaldo L, Hayashi M, Hoffmann S, Schechtman L, and Vanparys P. 2012. ECVAM prevalidation study on in vitro cell transformation assays: General outline and conclusions of the study. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **744** (1): 12–19.

Cox DR, Hinkley DV. 1974. Theoretical Statistics. Chapman and Hall.

ECHA, 2010. Practical guide 2: How to report weight of evidence. Available at

https://echa.europa.eu/documents/10162/.../pg_report_weight_of_evidence_en.pdf

EURL ECVAM. 2013. EURL ECVAM Recommendation on the cell transformation assays based on the Bhas 42 cell line. Available at https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-recommendations/files-bhas/EURL_ECVAM_Recommendation_Bhas-CTA_2013.pdf

EURL ECVAM. 2012. EURL ECVAM Recommendation on three Cell Transformation Assays using Syrian Hamster Embryo Cells (SHE) and the BALB/c 3T3 Mouse Fibroblast Cell Line for In Vitro Carcinogenicity Testing. Available at https://eurl-ecvam.jrc.ec.europa.eu/eurl-ecvam-recommendations/EURL-ECVAM -Recommendation.pdf

Forcella M, Callegaro G, Melchioretto P, Gribaldo L, Frattini M, Stefanini FM, Fusi P, Urani C. 2016. Cadmium-transformed cell in the *in vitro* Cell Transformation Assay reveal different proliferative behaviours and activated pathways. *Toxicol. in Vitro*. **36**: 71-80.

Fox J. 2003. Effect Displays in R for Generalised Linear Models. *J. Stat. Softw*. **8**(15): 1-27.

Goodman JE, Prueitt RL, Thakali S, and Oller AR. 2011. The nickel ion bioavailability model of the carcinogenic potential of nickel-containing substances in the lung. *Crit. Rev. Toxicol.* **41**(2): 142-174.

Haber LT, Erdreicht L, Diamond GL, Maier AM, Ratney R, Zhao Q, and Dourson ML. 2000. Hazard Identification and Dose Response of Inhaled Nickel-Soluble Salts. *Regul. Toxicol. Pharmacol*. **31**: 210-230.

Hoffmann S, Hothorn LA, Edler L, Kleensang A, Suzuki M, Phrakonkham P, and Gerhard D. 2012. Two new approaches to improve the analysis of BALB/c 3T3 cell transformation assay data. *Mutat. Res.* **744**: 36–41.

International Agency for Research on Cancer. 2012a. IARC Monograph: arsenic, metals, fibres, and dusts (volume 100 C). A review of human Carcinogens. Lyone, France.

International Agency for Research on Cancer. 2012b. IARC Monograph: chemical agents and related occupations (volume 100 F). A review of human Carcinogens. Lyone, France.

Keshava N. 2000. Tumorigenicity of morphologically distinct transformed *foci* induced by 3-methylcholanthrene in BALB/c-3T3 cells. *Mutat. Res.* **447** (2): 281-6.

Lu Y P, Lasne C, Chouroulinkov I. 1986. Use of an orthogonal design method to study two-stage chemical carcinogenesis in BALB/3T3 cells, Carcinogenesis **7**: 893–898.

Magaye R, Zhao J, Bowman L, Ding M. 2012. Genotoxicity and carcinogenicity of cobalt-, nickel- and copper-based nanoparticles. *Exp. Ther. Med*. **4**(4):551-561.

Maire MA, Rast C, Vasseur P. 2012. Photo catalogue for the classification of cell colonies in the Syrian hamster embryo (SHE) Cell Transformation Assay at pH 7.0. *Mutat. Res*. **744**(1): 97-110.

OECD 2016. Guidance Document on the In Vitro Bhas 42 Cell Transformation Assay.

Series on Testing & Assessment No. 231. (https://www.oecd.org/env/ehs/testing/ENV_JM_MONO(2016)1.pdf).

OECD 2015. Guidance Document on the In Vitro Syrian Hamster Embryo (SHE) Cell Transformation Assay. Series on Testing & Assessment No. 214. (http://www.oecd.org/env/ehs/testing/seriesontestingandassessmentpublicationsbynumber.htm)

OECD 2009. Test Guideline 451 – Carcinogenicity studies. OECD Guidelines for the Testing of Chemicals, OECD, Paris. (http://www.oecd-ilibrary.org/environment/test-no-451-carcinogenicity-studies_9789264071186-en.)

Pinheiro JC, and Bates DM. 2000. Chapter 5: Extending the Basic Linear Mixed-Effects Model. In *Mixed-Effects Models in S and S-PLUS*. Springer: New York; 201-267.

Pinheiro JC, Bates DM, DebRoy S, Sarkar D, and R Core Team. 2015. *nlme: Linear and Nonlinear Mixed Effects Models*, R package version 3.1-120. http://CRAN.R-project.org/package=nlme

Python Software Foundation, Python language reference, version 2.7. URL: https://www.python.org/.

Priya S, Nigam A, Bajpai P, Kumar S. 2013. Dysregulation of pathways involved in the processing of cancer and microenvironment information in MCA+TPA transformed C3H/10T1/2 cells. *In Vitro Cell. Dev. Biol. Anim.* **49**: 295-305.

Procaccianti C, Stefanini FM, Urani C. 2011. The Cell Transformation Assay: Toward a statistical classification of mixed and intermediate *foci* images. *ATLA*. **39**(1):23-36.

R Core Team. 2012. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL *http://www.R-project.org/*

Ridder GM, Stuard SB, Kerckaert GA, Cody DB, LeBoeuf LA, Isfort RJ. 1997. Computerized image analysis of morphologically transformed and nontransformed Syrian hamster embryo (SHE) cell colonies: application to objective SHE cell transformation assay scoring. *Carcinogenesis*. **18**: 1965-1972.

Rohrbeck A, Salinas G, Maaser K, Linge J, Salovaara S, Corvi R, Borlak J. 2010. Toxicogenomics applied to in vitro carcinogenicity testing with Balb/c 3T3 cells revealed a gene signature predictive of chemical carcinogens. *Toxicol. Sci.* **118**: 31-41.

Sakai A, Sasaki K, Hayashi K, Muramatsu D, Arai S, Endou N, Kuroda S, Poth A, Bohnenberger S, Kunkelmann T, Asakura M, Hirose H, Ishii N, Mizuhashi F, Kasamoto S, Nagai M, Pant K, Bruce SW, Sly JE, Yamazaki S, Umeda M, Tanaka N. 2011. An international validation study of a Bhas 42 cell transformation assay for the prediction of chemical carcinogenicity.

Mutat. Res. **725**(1-2):57-77.

Sasaki K, Umeda M, Sakai A, Yamazaki S, Tanaka N. 2015. Transformation assay in Bhas 42 cells: a model using initiated cells to study mechanisms of carcinogenesis and predict carcinogenic potential of chemicals. *J. Environ. Sci. Health C. Environ. Carcinog. Ecotoxicol. Rev*. **33**:1–35.

Sasaki K, Bohnenberger S, Hayashi K, Kunkelmann T, Muramatsu D, Poth A, Sakai A, Salovaara S, Tanaka N, Thomas BC, and Umeda M. 2012a. Photo catalogue for the classification of *foci* in the BALB/c 3T3 cell transformation assay. *Mutat. Res*. **744**: 42–53.

Sasaki K, Bohnenberger S, Hayashi K, Kunkelmann T, Muramatsu D, Phrakonkham P, Poth A, Sakai A, Salovaara S, Tanaka N, Thomas BC, and Umeda M. 2012b. Recommended protocol for the BALB/c 3T3 cell transformation assay. *Mutat. Res*. **744**: 30–35.

Smets LA. 1980. Cell transformation as a model for tumor induction and neoplastic growth. *Biochim. Biophys. Acta.* **605** (1): 93-111.

Stefanini, FM. 2013. Comment: Bayesian network integrated testing strategy and beyond. Altex 30, 386-390.

Sun,H, Shamy M, and Costa M. 2013. Nickel and Epigenetic Gene Silencing. *Genes (Basel)* **4**, 583–595.

Tanaka N, Bohnenberger S, Kunkelmann T, Munaro B, Ponti J, Poth A, Sabbioni E, Sakai A, Salovaara S, Sasaki K, Thomas BC, Umeda M. 2012. Prevalidation study of the BALB/c 3T3 cell transformation assay for assessment of carcinogenic potential of chemicals. *Mutat. Res. Genet. Toxicol. Environ. Mutagen*. **744**(1): 20-29.

Urani C, Corvi R, Callegaro G, and Stefanini FM. 2013. Objective scoring of transformed foci in BALB/c 3T3 cell transformation assay by statistical image descriptors. *Toxicol. in Vitro* **27**: 1905–1912.

Urani C, Stefanini FM, Bussinelli L, Melchioretto P, Crosta GF. 2009. Image analysis and automatic classification of transformed *foci*. *J. Microsc*. 234: 269-279.

Walsh MJ, Bruce SW, Pant K, Carmichael PL, Scott AD and Martin FL. 2009. Discrimination of a transformation phenotype in Syrian golden hamster embryo (SHE) cells using ATR-FTIR spectroscopy. *Toxicology*. **258**(1):33-8.

TABLES

Table 1. The number of dishes and *foci* included in the final database following the segmentation process are shown for each $NiCl_2$ concentration. The initial number of dishes available for each concentration was 10.

Dose (μM)	# dishes containing Type IIII *foci* at the end of CTA	# dishes considered after segmentation	total # of *foci* (III) before segmentation	# *foci* (III) after segmentation
250	7	4	10	6
275	10	10	33	23
300	10	10	73	40
350	10	10	137	78
400	9	9	29	18
Total	46	43	282	165

Table 2. The number of dishes and *foci* included in the final database following the segmentation process are shown for each B[a]P concentration. The initial number of dishes available for each concentration was 10.

Dose (μg/ml)	# dishes containing Type IIII foci at the end of CTA	# dishes considered after segmentation	total # of foci (III) before segmentation	# foci (III) after segmentation
0.0005	4	4	6	4
0.001	2	2	2	2
0.005	7	7	15	13
0.05	10	10	70	38
0.125	10	10	92	44
0.625	10	10	111	73
3.125	10	10	74	41
15	8	7	12	9
Total	46	43	382	224

Table 3. Point estimates from the $NiCl_2$ dataset. Expected value functions of each univariate conditional distribution for the final models including $NiCl_2$ concentration before testing the null hypothesis of no effect by means of likelihood ratio tests. The estimated value of intraclass correlation parameter is indicated as $\hat{\rho}$. The variance was modelled as a function of the rescaled concentration $\tilde{x}_C$. In particular, the estimated conditional variance function in all models is defined by the following equation: $\sigma_y^2 = \sigma^2(\tilde{x}_C)^{2\delta}$, where y is the descriptor considered as a response variable.

Note that $\hat{Y}_{ED} \equiv E[Y_{ED}|\, y_{MD}, y_{BD}, x_{PD}, x_C, \theta_{ED}]$.

Expected value	Variance function	Correlation
N1: $\hat{Y}_{ED} = 51993.62 - 141843.64y_{MD} + 97223.63y_{MD}^2 + + 1437.82y_{BD} - 108.02x_C - 1848.83y_{MD}y_{BD}$	$\hat{\sigma} = 242.08$ $\hat{\delta} = 0.42$	$\hat{\rho} = 0.15$
N2: $\hat{Y}_{MD} = 0.72 + 0.013y_{BD} - 0.13x_C + 0.13x_C^2$	$\hat{\sigma} = 0.024$ $\hat{\delta} = -0.16$	$\hat{\rho} = 0.01$
N3: $\hat{Y}_{BD} = 1.25 + 1.28x_C$	$\hat{\sigma} = 0.94$ $\hat{\delta} = 0.16$	$\hat{\rho} = 0.006$

Table 4. Point estimates from the B[a]P dataset. Expected value functions of each univariate conditional distribution for the final models including B[a]P concentration before testing the null hypothesis of no effect by means of likelihood ratio tests. The estimated value of intraclass correlation parameter is indicated as $\hat{\rho}$. The variance was modelled as a function of the rescaled concentration $\tilde{x}_C$. In particular, the estimated conditional variance function in all models is defined by the following equation: $\sigma_y^2 = \sigma^2(\tilde{x}_C)^{2\delta}$, where y is the descriptor considered as a response variable.

Note that $\hat{Y}_{ED} \equiv E[Y_{ED}|\, y_{MD}, y_{BD}, x_{PD}, x_C, \theta_{ED}]$.

Expected value	Variance function	Correlation
B1: $\hat{Y_{ED}} = 3232.74 - 4405.34y_{MD} + 247.31y_{BD} - 38.68y_{BD}^2 - 0.81x_C$	$\hat{\sigma} = 228.70$ $\hat{\delta} = -0.06$	$\hat{\rho} = 0.15$
B2: $\hat{Y}_{MD} = 0.67 + 0.003y_{BD} + 0.001x_C$	$\hat{\sigma} = 0.016$ $\hat{\delta} = 0.03$	$\hat{\rho} = 0.06$
B3: $\hat{Y}_{BD} = 1.60 + 0.08x_C$	$\hat{\sigma} = 0.72$ $\hat{\delta} = -0.06$	$\hat{\rho} = 0.16$

FIGURE CAPTIONS

Figure 1. Experimental workflow. Main steps of the workflow: following the CTA experiments (1), *foci* images were acquired (2) and from each image three statistical descriptors were selected to represent distinct morphological features used in visual scoring (3). The *foci* datasets of the two carcinogens were analyzed by factorizing the joint probability distribution into three conditional distributions (5a), each one subsequently modeled in the class of extended linear mixed-effects models (5b). The final models including the concentration were selected according to the BIC value, then they were exploited to test the effect of carcinogen concentration on each image descriptors by likelihood ratio tests (5c).

Figure 2. Boxplots of the number of *foci*. Boxplots of the number of Type III *foci* per dish given the carcinogen concentration. In A the number of *foci* from $NiCl_2$ dataset are shown, while in B the number of *foci* from B[a]P dataset are displayed. Empty circles in B represent candidate outliers.

Figure 3. The dependence relationships between descriptors and concentration. The descriptors (ED, BD, MD) are displayed in blue circles, while nodes for concentration of a carcinogen (C) and Petri dish, both conditioning variables, are shown in yellow. Details of dependencies are provided in *section 2.4.1.*

Figure 4. Plot of expected value functions for the Equivalent Diameter (ED) given the carcinogen concentration. In C and H, expected value functions of ED in models N1 ($NiCl_2$) and B1 (B[a]P) are plotted given the concentration. The grey shaded area represents pointwise confidence limits, on the scale of the linear predictor. Examples of *foci s*howing ED values that can be found in the extreme regions of the graph are displayed in the panels A, B, D and E for model N1, and in F, G, I and J for model B1. The boxes on the x axis represent conditioning values of the carcinogen concentration.

Figure 5. Plot of expected value functions of Median (MD) given the carcinogen concentration. In B and F expected values of MD in model N2 ($NiCl_2$) and B2 (B[a]P), respectively are plotted given the concentration. The grey shaded area represents pointwise confidence limits, on the scale of the linear predictor. Examples of *foci* showing MD values corresponding to the area inside the blue circles are displayed in panels A, C and D, while examples of *foci* showing MD values corresponding to the area inside the red circles are displayed in panels E and G. The boxes on the x axis represent conditioning values of carcinogen concentration.

Figure 6. Plot of expected value functions of Boundary Index (BD) given the carcinogen concentration. In B and E, expected values of BD in model N3 ($NiCl_2$) and B3 (B[a]P), respectively are plotted given the concentration. The grey shaded area represents pointwise confidence limits, on the scale of the linear predictor. Examples of *foci* showing BD values corresponding to the area inside the blue circles are displayed in panels A and C, while D and F show *foci* having BD values corresponding to the area inside the red circles. The boxes on the x axis represent conditioning values of carcinogen concentration.

SUPPORTING INFORMATION

Figure S1. Boxplots of the number of *foci* before performing segmentation. Boxplots of the number of Type III *foci* per dish (y axes) before segmentation process and carcinogen concentrations (x axes). In A *foci* images from $NiCl_2$ database are considered, while in B *foci* images from B[a]P database are included. Empty circles represent candidate outliers.

Figure S2. Diagnostics and additional plots of models of $NiCl_2$ dataset. Graphs are grouped according to model, by column: Model N1 (ED as the response variable), Model N2 (MD as the response variable) and Model N3 (BD as the response variable). In the first two rows diagnostic graphs are shown, quantile-quantile plots (A, B, C) and plot of fitted values versus residuals (D, E, F), respectively. In G and I are plotted fitted values (Model N1) of the descriptor ED compared, respectively, to the linear predictor MD and to the linear predictor BD. The bars on the x axis represent MD and BD actual values, in plot C and D, respectively. In H are plotted fitted values (Model N2) of the descriptor MD compared to the linear predictor MD (bars on the x axis represent BD actual values).

Figure S3. Diagnostics and additional plots of models of B[a]P dataset. Graphs are grouped according to model, by column: Model B1 (ED as the response variable), Model B2 (MD as the response variable) and Model B3 (BD as the response variable). In the first two rows diagnostic graphs are shown, quantile-quantile plots (A, B, C) and plot of fitted values versus residuals (D, E, F), respectively. In G and I are plotted fitted values (Model B1) of the descriptor ED compared, respectively, to the linear predictor MD and to the linear predictor BD. The bars on the x axis represent MD and BD actual values, in plot C and D, respectively. In H are plotted fitted values (Model B2) of the descriptor MD compared to the linear predictor MD (bars on the x axis represent

BD actual values).

Report S4. Supplementary statistical details and R code.

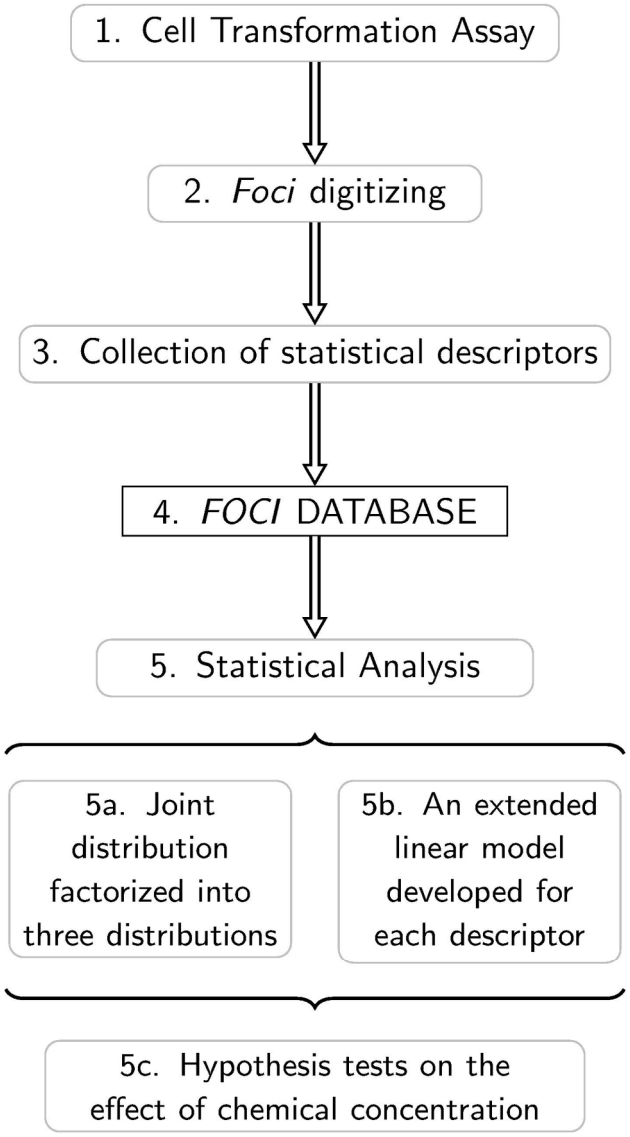


Figure 1. Experimental workflow. Main steps of the workflow: following the CTA experiments (1), *foci* images were acquired (2) and from each image three statistical descriptors were selected to represent distinct morphological features used in visual scoring (3). The *foci* datasets of the two carcinogens were analyzed by factorizing the joint probability distribution into three conditional distributions (5a), each one subsequently modeled in the class of extended linear mixed-effects models (5b). The final models including the concentration were selected according to the BIC value, then they were exploited to test the effect of carcinogen concentration on each image descriptors by likelihood ratio tests (5c).

116x225mm (300 x 300 DPI)

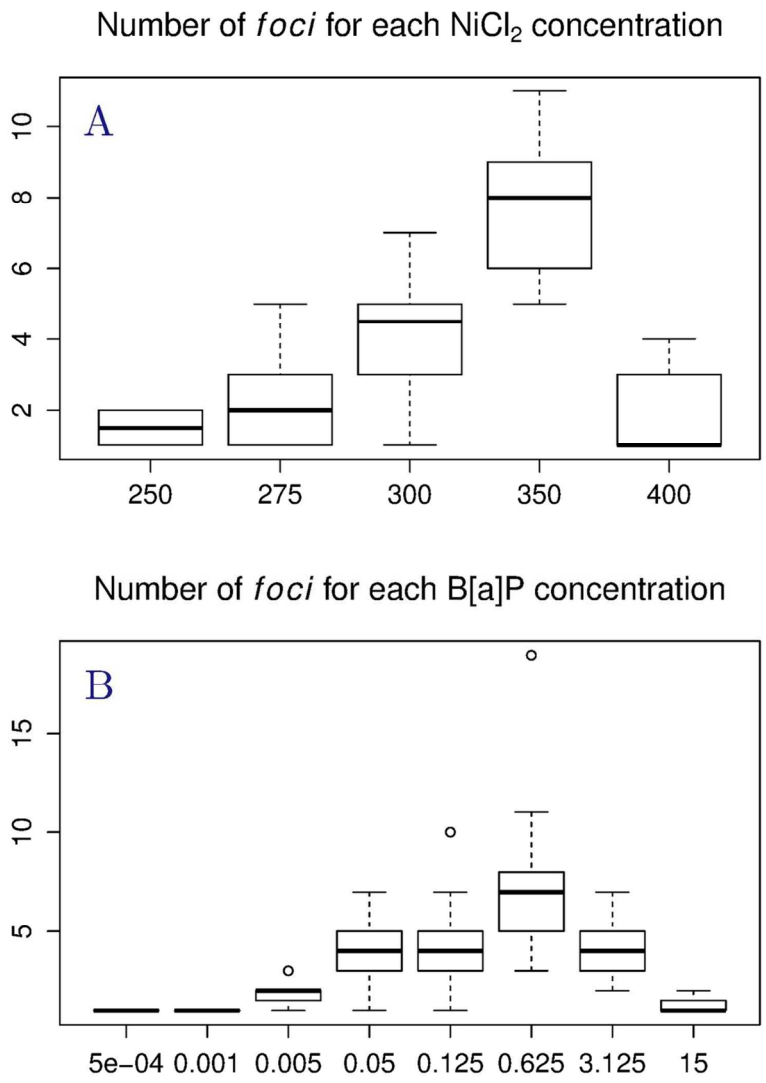


Figure 2. Boxplots of the number of *foci*. Boxplots of the number of Type III *foci* per dish given the carcinogen concentration. In A the number of foci from $NiCl_2$ dataset are shown, while in B the number of *foci* from B[a]P dataset are displayed. Empty circles in B represent candidate outliers.

98x134mm (300 x 300 DPI)

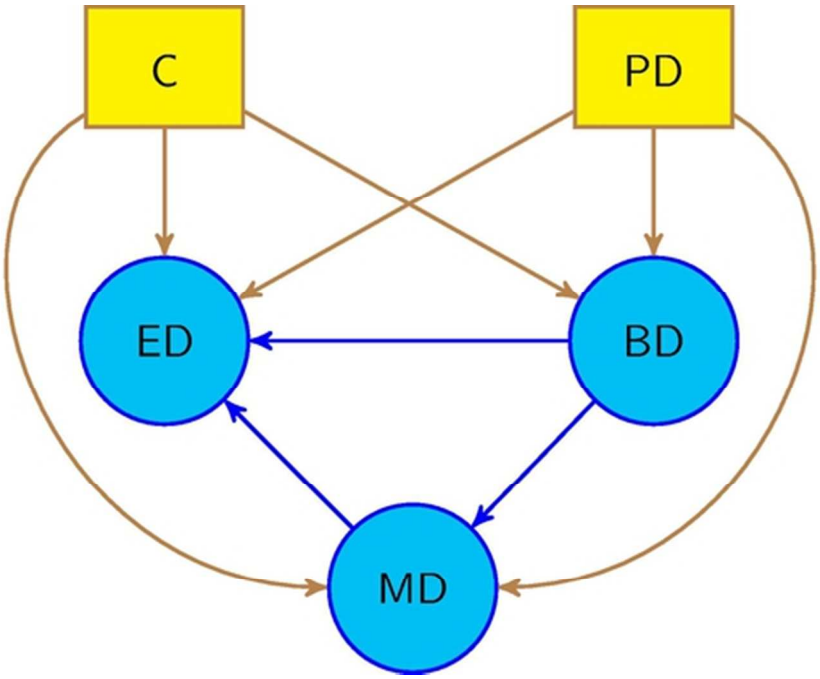


Figure 3. The dependence relationships between descriptors and concentration. The descriptors (ED, BD, MD) are displayed in blue circles, while nodes for concentration of a carcinogen (C) and Petri dish, both conditioning variables, are shown in yellow. Details of dependencies are provided in *section 3.1*.

52x35mm (300 x 300 DPI)

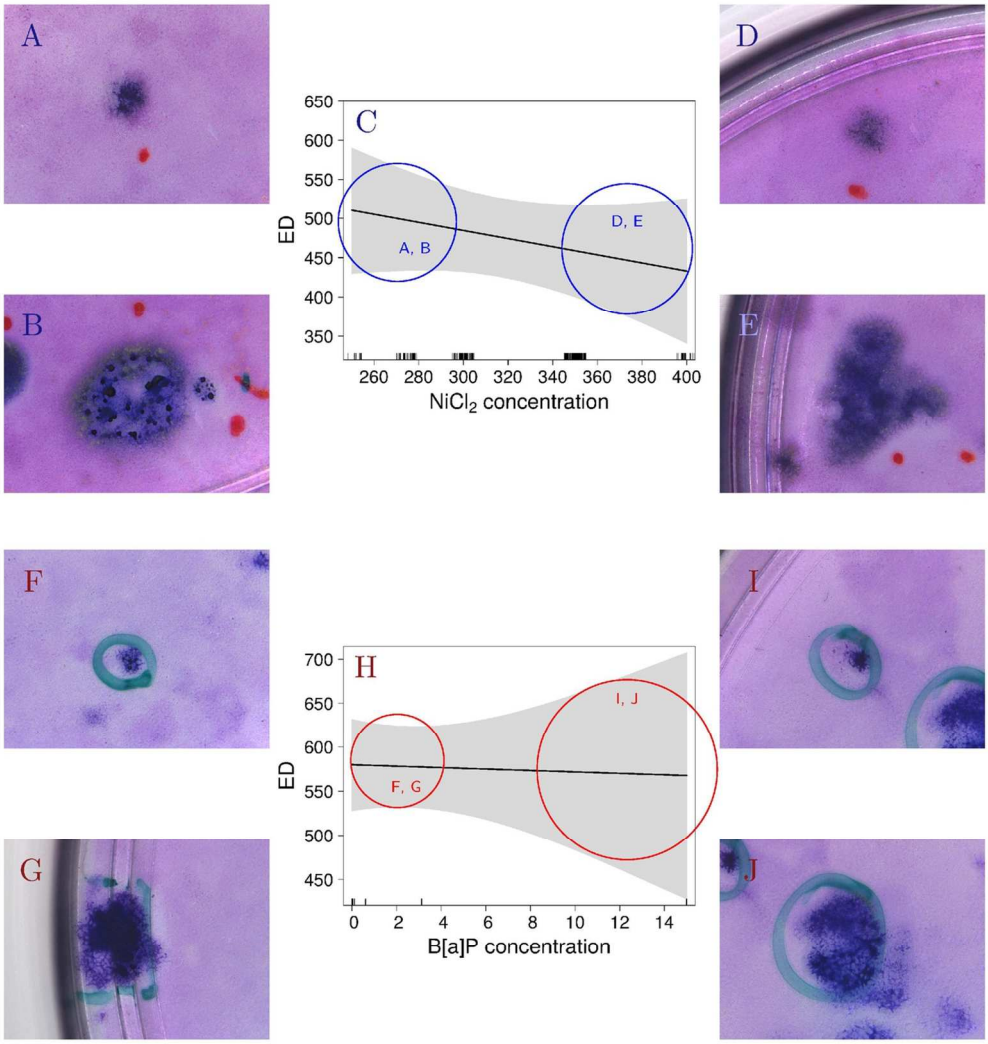


Figure 4. Plot of expected value functions for the Equivalent Diameter (ED) given the carcinogen concentration. In C and H, expected value functions of ED in models N1 ($NiCl_2$) and B1 (B[a]P) are plotted given the concentration. The grey shaded area represents pointwise confidence limits, on the scale of the linear predictor. Examples of *foci* showing ED values that can be found in the extreme regions of the graph are displayed in the panels A, B, D and E for model N1, and in F, G, I and J for model B1. The boxes on the x axis represent conditioning values of the carcinogen concentration.

119x125mm (300 x 300 DPI)

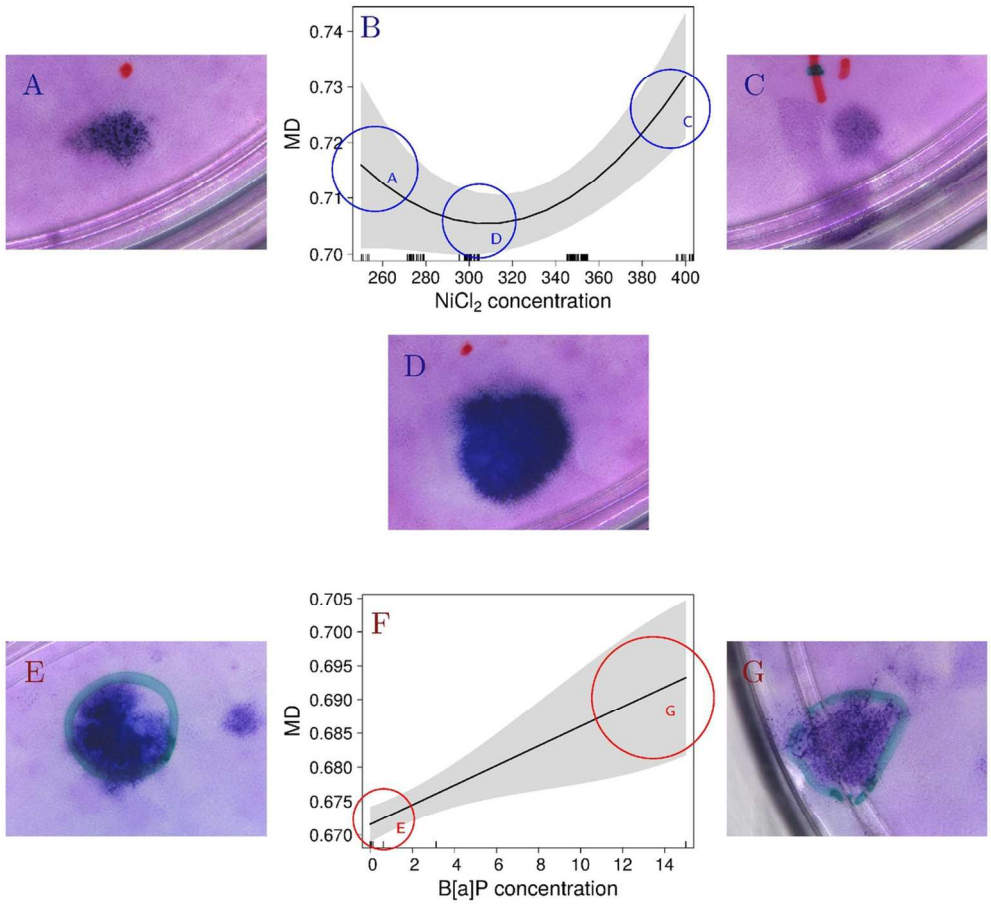


Figure 5. Plot of expected value functions of Median (MD) given the carcinogen concentration. In B and F expected values of MD in model N2 ($NiCl_2$) and B2 (B[a]P), respectively are plotted given the concentration. The grey shaded area represents pointwise confidence limits, on the scale of the linear predictor. Examples of *foci* showing MD values corresponding to the area inside the blue circles are displayed in panels A, C and D, while examples of *foci* showing MD values corresponding to the area inside the red circles are displayed in panels E and G. The boxes on the x axis represent conditioning values of carcinogen concentration.

109x103mm (300 x 300 DPI)

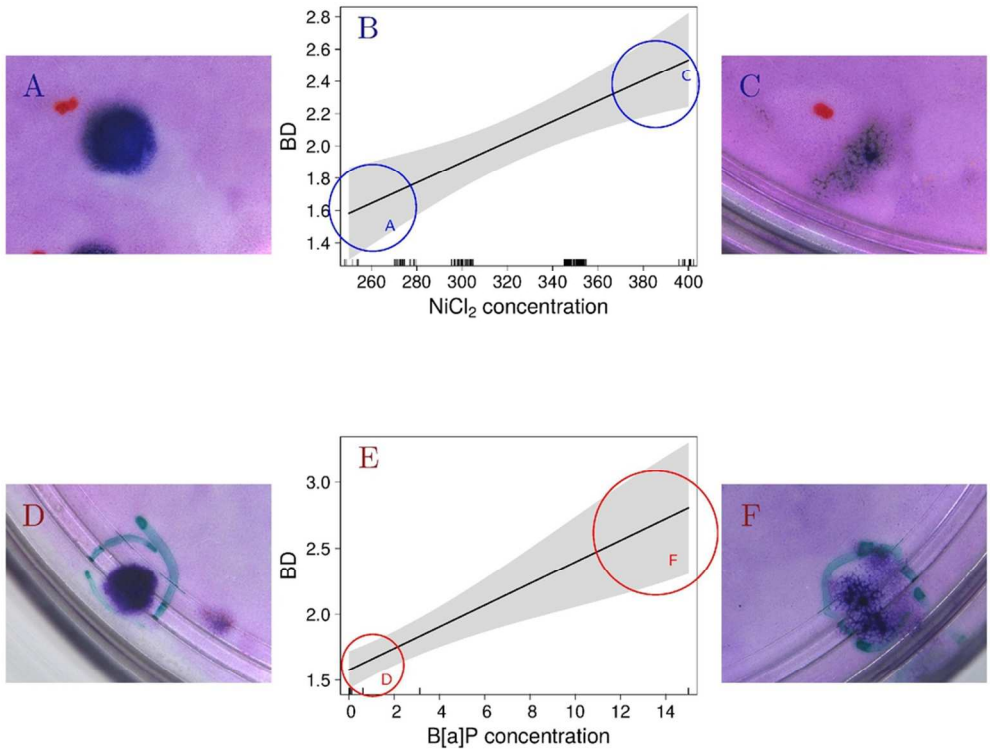
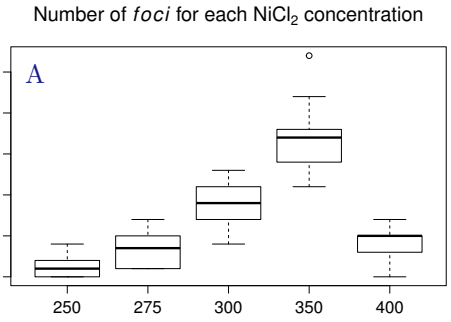
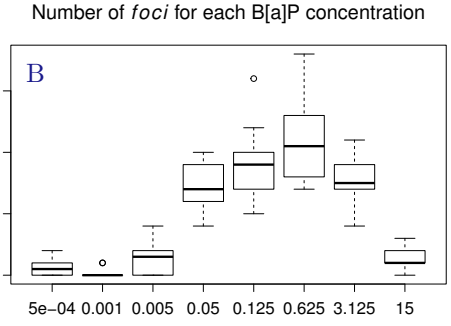
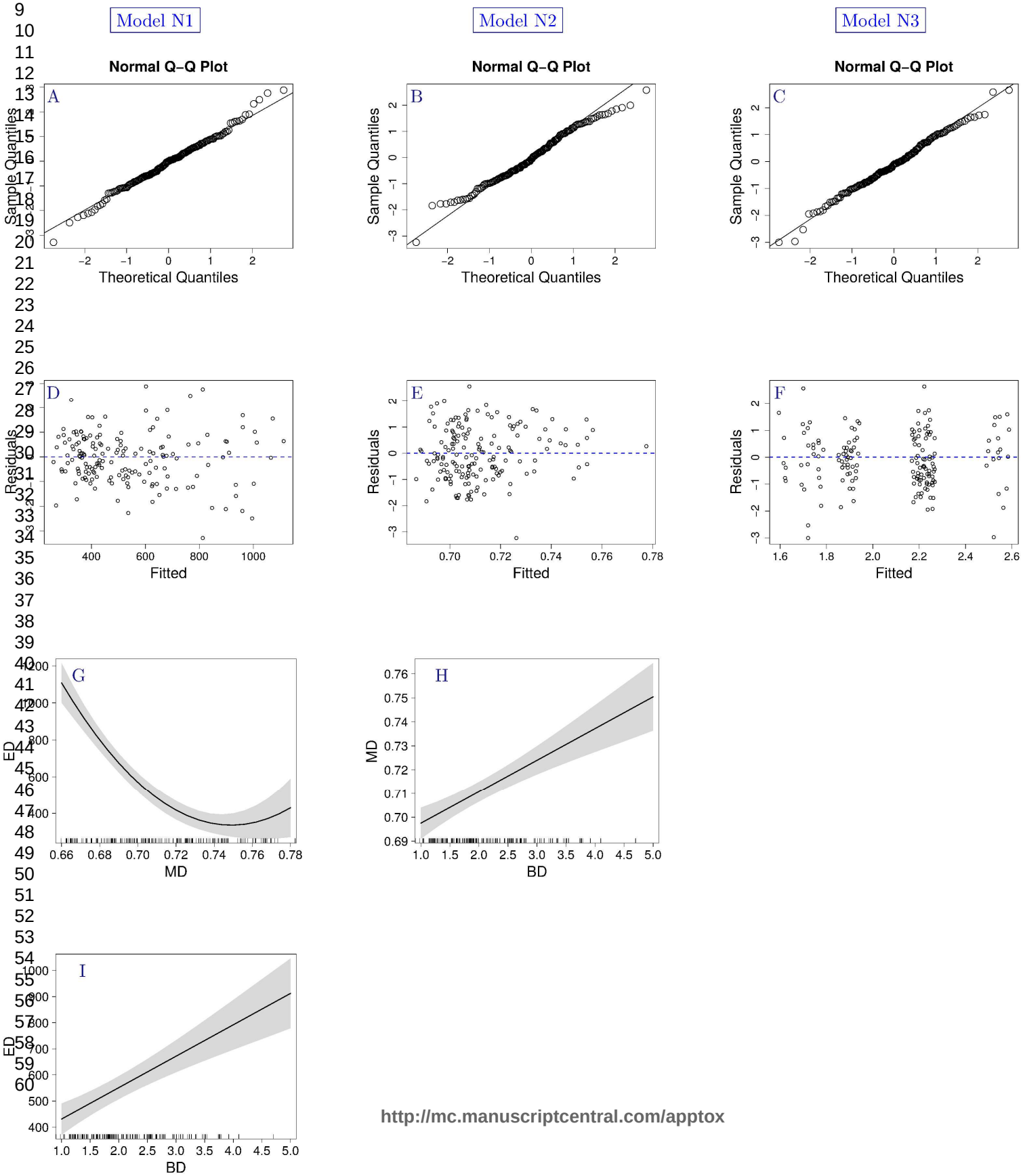


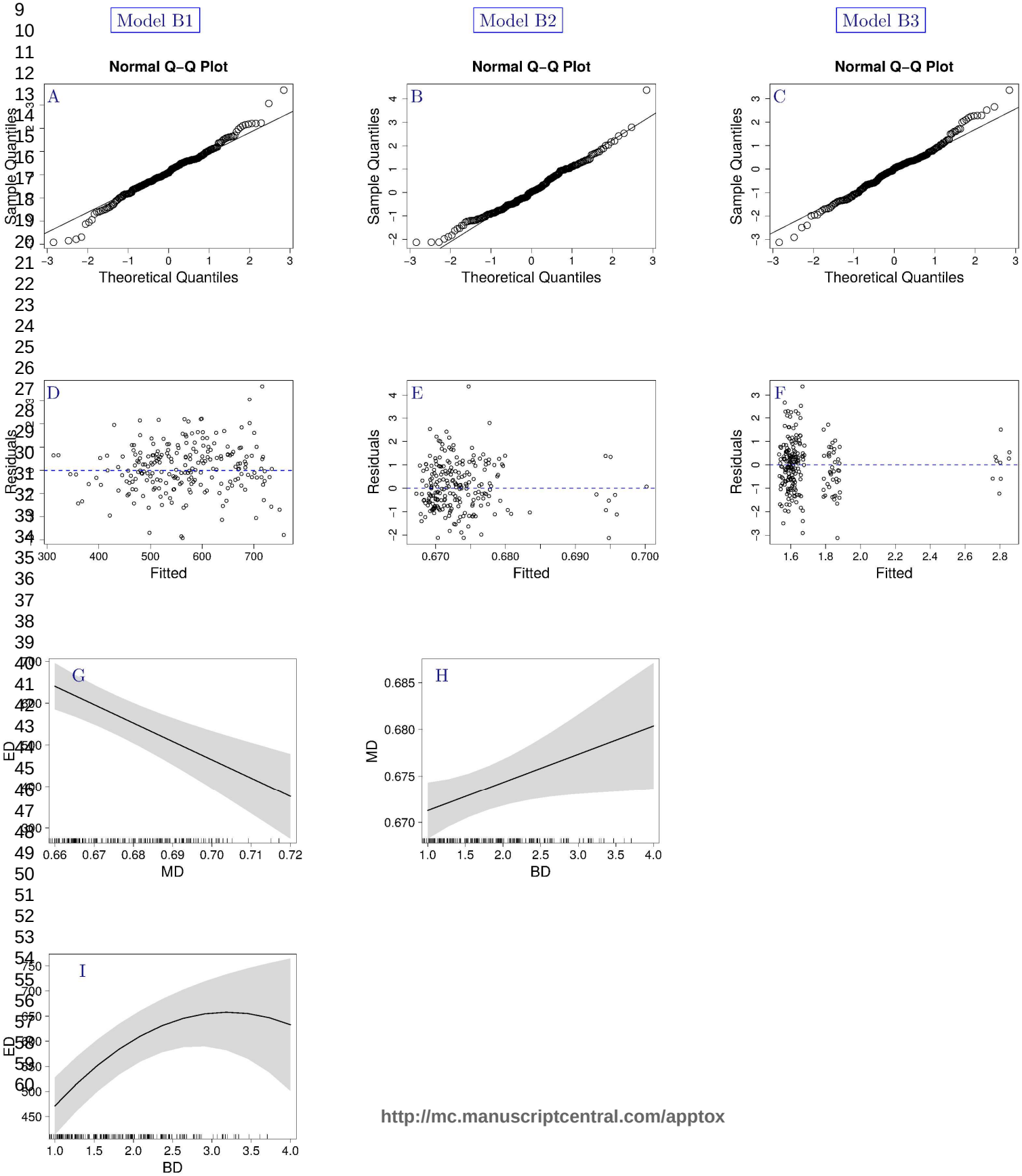
Figure 6. Plot of expected value functions of Boundary Index (BD) given the carcinogen concentration. In B and E, expected values of BD in model N3 ($NiCl_2$) and B3 (B[a]P), respectively are plotted given the concentration. The grey shaded area represents pointwise confidence limits, on the scale of the linear predictor. Examples of *foci* showing BD values corresponding to the area inside the blue circles are displayed in panels A and C, while D and F show *foci* having BD values corresponding to the area inside the red circles. The boxes on the x axis represent conditioning values of carcinogen concentration.

90x72mm (300 x 300 DPI)






Model N1
Model N2
Model N3
Normal Q–Q Plot
Normal Q–Q Plot
Normal Q–Q Plot
A
B
C
Sample Quantiles
Theoretical Quantiles
D
E
F
Residuals
Fitted
G
H
I
ED
MD
BD


Model B1
Model B2
Model B3
Normal Q–Q Plot
Sample Quantiles
Theoretical Quantiles
A
B
C
Residuals
Fitted
D
E
F
ED
MD
G
H
BD
I

Supporting Material S4

This document is provided as supporting material for the paper "Relationship between increasing concentrations of a genotoxic and a non-genotoxic carcinogen and morphology in the Cell Transformation Assay as described by statistical image descriptors".

In the following sections R code for the statistical analysis performed in the main article is detailed, as well as supplementary materials of the presented models.

First, the following packages have to be loaded:

```
library(nlme)
library(car)
library(nlme)
library(effects)
```

First, models for each descriptor as a response variable are fitted with generalized least squares accounting for unequal variance, and including a correlation structure (object of class).

```
mod.1N <- gls(equivalentDiameter~median
              +I(median^2)
              +boundaryIndex
              +median:boundaryIndex
              +dose,
              data=workDF,
              correlation = corCompSymm(0.1, form = ~ 1 | plate),
              method='ML')
mod.1N <- update(mod.1N, weights = varPower(form= ~ dose))

mod.2N <- gls(median~boundaryIndex
                +dose
                +I(dose^2),
                data=workDF,
                correlation = corCompSymm(0.1, form = ~ 1 | plate),
                method='ML')
mod.2N <- update(mod.2N, weights = varPower(form = ~dose))

mod.3N <- gls(boundaryIndex~dose,
                data=workDF,
                correlation = corCompSymm(0.1, form= ~1 |plate),
                method='ML')
mod.3N <- update(mod.3N, weights = varPower(form = ~ dose))
```

Then contribution of concentration on the expected value of descriptors was tested with reducted models:

```r
mod.1N_r <- update(mod.1N, ~.-dose)
anova(mod.1N, mod.1N_r )

mod.2N_r <- update(mod.2N, ~.-dose - I(dose^2))
anova(mod.2N, mod.2N_r )

mod.3N_r <- update(mod.3N, ~.-dose)
anova(mod.3N, mod.3N_r )
```

Diagnostic graphs were created with the following code (mod.1N shown as example), and are provided in Figure S2.

```r
myres <- residuals(mod.1N, type='normalized')
qqnorm(myres)
qqline(myres)

plot(myres~jitter(fitted(mod.1N), amount=.05), data=workDF)
abline(h=0, col=4, lwd=2, lty=2)
```

We here provide, as supporting materials, confidence intervals for estimated parameters and variance function for each fitted model:

```r
intervals(mod.1N)
```

```
## Approximate 95% confidence intervals
##
##  Coefficients:
##                              lower         est.        upper
## (Intercept)            34516.6646   51993.6151  69470.56561
## median               -192036.6855 -141843.6447 -91650.60389
## I(median^2)            61201.4529   97223.6306 133245.80831
## boundaryIndex            609.1660    1437.8238   2266.48156
## doseNORM                -290.7646    -108.0184     74.72777
## median:boundaryIndex   -2990.1239   -1848.8325   -707.54106
## attr(,"label")
## [1] "Coefficients:"
##
##  Correlation structure:
##          lower      est.     upper
## Rho 0.03874005 0.1541243 0.3232199
## attr(,"label")
## [1] "Correlation structure:"
##
##  Variance function:
##            lower      est.     upper
## power 0.01743034 0.4164543 0.8154782
## attr(,"label")
## [1] "Variance function:"
##
##  Residual standard error:
```

```
##    lower     est.    upper
## 194.0061 242.0850 302.0790

intervals(mod.2N)

## Approximate 95% confidence intervals
##
##  Coefficients:
##                      lower        est.       upper
## (Intercept)    0.674492416  0.71539245  0.75629247
## boundaryIndex  0.008647398  0.01326386  0.01788032
## doseNORM      -0.262721928 -0.13779164 -0.01286134
## I(doseNORM^2)  0.032246221  0.12622566  0.22020509
## attr(,"label")
## [1] "Coefficients:"
##
##  Correlation structure:
##           lower        est.     upper
## Rho -0.04937754 0.009573488 0.1226112
## attr(,"label")
## [1] "Correlation structure:"
##
##  Variance function:
##            lower      est.     upper
## power -0.5023479 -0.159366 0.1836158
## attr(,"label")
## [1] "Variance function:"
##
##  Residual standard error:
##      lower       est.      upper
## 0.01972629 0.02402997 0.02927258

intervals(mod.3N)

## Approximate 95% confidence intervals
##
##  Coefficients:
##                 lower     est.    upper
## (Intercept) 0.8066259 1.249700 1.692775
## doseNORM    0.6118222 1.282781 1.953740
## attr(,"label")
## [1] "Coefficients:"
##
##  Correlation structure:
##          lower       est.     upper
## Rho -0.0687326 0.00572599 0.2066248
## attr(,"label")
## [1] "Correlation structure:"
##
##  Variance function:
##            lower      est.     upper
```

```
## power -0.1349208 0.1571835 0.4492878
## attr(,"label")
## [1] "Variance function:"
##
##  Residual standard error:
##     lower      est.     upper
## 0.7873429 0.9404370 1.1232993
```

Plots of fitted values of each descriptor as response variable for each model and concentration variable was obtained with the following code (mod.1N shown as example):

```
eff1N_dose <- Effect('dose', mod.1N,
                     xlevels = list(dose = unique(workDF$dose)),
                     vcov.=vcov, se = TRUE, confidence.level = 0.95,
                     transformation = NULL,
                     typical = mean)
plot(eff1N_dose)
```

First, models for each descriptor as a response variable are fitted with generalized least squares accounting for unequal variance, and including a correlation structure (object of class).

```
mod.1B <- gls(equivalentDiameter~median
              +boundaryIndex
              +I(boundaryIndex^2)
              +dose,
              data=workDF,
              correlation = corCompSymm(0.1, form = ~ 1 | plate),
              method='ML')
mod.1B <- update(mod.1B, weights = varPower(form= ~ dose))

mod.2B <- gls(median~boundaryIndex
             +dose,
             data=workDF,
             correlation = corCompSymm(0.1, form = ~ 1 | plate),
             method='ML')
mod.2B <- update(mod.2B, weights = varPower(form =~dose))

mod.3B <- gls(boundaryIndex~dose,
              data=workDF,
              correlation = corCompSymm(0.1, form= ~1 |plate),
              method='ML')
mod.3B <- update(mod.3B, weights = varPower(form = ~ dose))
```

Then contribution of concentration was tested with reducted models:

```
mod.1B_r <- update(mod.1B, ~.-dose)
anova(mod.1B, mod.1B_r)

mod.2B_r <- update(mod.2B, ~.-dose)
anova(mod.2B, mod.2B_r)
```

```
mod.3B_r <- update(mod.3B, ~.-dose)
anova(mod.3B, mod.3B_r)
```

Diagnostic graphs were created with the following code (mod.1B shown as example), and are provided in Figure S3.

```
myres <- residuals(mod.1B, type='normalized')
qqnorm(myres, main='q-q plot mod.1B')
qqline(myres)

plot(myres~jitter(workDF$doseSTD, amount=.05), data=workDF)
abline(h=0, col=4, lwd=2, lty=2)
```

We here provide, as supporting materials, confidence intervals for estimated parameters and variance function for each fitted model:

```
intervals(mod.1B)

## Approximate 95% confidence intervals
##
##  Coefficients:
##                           lower          est.         upper
## (Intercept)         1898.17101  3232.7439192  4567.316824
## median             -6406.81328 -4405.3403410 -2403.867401
## boundaryIndex         93.19703   247.3053228   401.413611
## I(boundaryIndex^2)   -72.20875   -38.6749767    -5.141206
## dose                 -11.28108    -0.8103561     9.660365
## attr(,"label")
## [1] "Coefficients:"
##
##  Correlation structure:
##          lower      est.     upper
## Rho 0.03737678 0.1450291 0.3018267
## attr(,"label")
## [1] "Correlation structure:"
##
##  Variance function:
##            lower        est.       upper
## power -0.1072438 -0.05933182 -0.01141989
## attr(,"label")
## [1] "Variance function:"
##
##  Residual standard error:
##    lower     est.    upper
## 203.7128 228.6996 256.7512

intervals(mod.2B)

## Approximate 95% confidence intervals
##
```

```
##  Coefficients:
##                       lower        est.       upper
## (Intercept)   0.6611820055 0.666215044 0.671248082
## boundaryIndex 0.0002813363 0.003027738 0.005774139
## dose          0.0006216444 0.001445237 0.002268830
## attr(,"label")
## [1] "Coefficients:"
##
##  Correlation structure:
##           lower       est.     upper
## Rho -0.007080117 0.06425425 0.1753773
## attr(,"label")
## [1] "Correlation structure:"
##
##  Variance function:
##            lower       est.      upper
## power -0.01543731 0.03072261 0.07688253
## attr(,"label")
## [1] "Variance function:"
##
##  Residual standard error:
##      lower       est.      upper
## 0.01459902 0.01628850 0.01817350

intervals(mod.3B)

## Approximate 95% confidence intervals
##
##  Coefficients:
##                  lower       est.     upper
## (Intercept) 1.44733940 1.58828135 1.7292233
## dose        0.04595192 0.07733259 0.1087133
## attr(,"label")
## [1] "Coefficients:"
##
##  Correlation structure:
##          lower      est.     upper
## Rho 0.05466433 0.1577961 0.3005583
## attr(,"label")
## [1] "Correlation structure:"
##
##  Variance function:
##            lower        est.      upper
## power -0.2263946 -0.06441809 0.09755844
## attr(,"label")
## [1] "Variance function:"
##
##  Residual standard error:
##     lower      est.     upper
## 0.6090554 0.7188179 0.8483614
```

Plots of fitted values of each descriptor as response variable for each model and concentration variable was obtained with the following code (mod.1B shown as example):

```
eff1B_dose <- Effect('dose', mod.1B,
                     xlevels = list(dose = unique(workDF$dose)),
                     vcov.=vcov, se = TRUE, confidence.level = 0.95,
                     transformation = NULL,
                     typical = mean)
plot(eff1B_dose)
```