Recommendations, evaluation and validation of a semi-automated, fluorescent-based scoring protocol for micronucleus testing in human cells

Anna L. Seager*, Ume-Kulsoom Shah, Katja Brüsehafer, John Wills, Bella Manshian, Katherine E. Chapman, Adam D. Thomas, Andrew D. Scott1, Ann T. Doherty2, Shareen H. Doak, George E. Johnson and Gareth J. S. Jenkins

DNA Damage Research Group, Institute of Life Science, College of Medicine, Swansea University, Swansea SA2 8PP, UK, 1Safety and Environmental Assurance Centre (SEAC), Unilever, Colworth Science Park, Sharnbrook, Bedfordshire MK44 ILQ, UK and 2Genetic Toxicology, Safety Assessment, AstraZeneca R&D, Alderley Park, Macclesfield, Cheshire SK10 4TG, UK

*To whom correspondence should be addressed. Institute of Life Science, College of Medicine, Swansea University, Singleton Park, Swansea SA2 8PP, UK. Tel: +44 179 229 5056; Fax: +44 179 260 2147; Email: a.l.seager@swansea.ac.uk

Received on March 28, 2013; revised on November 14, 2013; accepted on November 19, 2013

Micronucleus (MN) induction is an established cytogenetic end point for evaluating structural and numerical chromosomal alterations in genotoxicity testing. A semi-automated scoring protocol for the assessment of MN preparations from human cell lines and a 3D skin cell model has been developed and validated. Following exposure to a range of test agents, slides were stained with 4′-6-diamidino-2-phenylindole (DAPI) and scanned by use of the MicroNuc module of metafer 4, after the development of a modified classifier for selecting MN in binucleate cells. A common difficulty observed with automated systems is an artefactual output of high false positives, in the case of the metafer system this is mainly due to the loss of cytoplasmic boundaries during slide preparation. Slide quality is paramount to obtain accurate results. We show here that to avoid elevated artefactual-positive MN outputs, diffuse cell density and low-intensity nuclear staining are critical. Comparisons between visual (Giemsa stained) and automated (DAPI stained) MN frequencies and dose-response curves were highly correlated ($R^2 = 0.70$ for hydrogen peroxide, $R^2 = 0.98$ for menadione, $R^2 = 0.99$ for mitomycin C, $R^2 = 0.89$ for potassium bromate and $R^2 = 0.68$ for quantum dots), indicating the system is adequate to produce biologically relevant and reliable results. Metafer offers many advantages over conventional scoring including increased output and statistical power, and reduced scoring subjectivity, labour and costs. Further, the metafer system is easily adaptable for use with a range of different cells, both suspension and adherent human cell lines. Awareness of the points raised here reduces the automatic positive errors flagged and drastically reduces slide scoring time, making metafer an ideal candidate for genotoxic biomonitoring and population studies and regulatory genotoxic testing.

Introduction

The importance of in vitro methods in toxicological assessment of chemicals/ pharmaceuticals is ever increasing, particularly following recent legislation including the Seventh Amendment to the Cosmetics Directive in 2009, which bans the use of in vivo assays for genotoxic assessment of cosmetic ingredients in the European Union. Other directives also promote alternatives to in vivo testing strategies in toxicology such as the National Centre for the Replacement, Refinement and Reduction of Animals in Research mission, which is tasked by UK Government with supporting the UK science base through the application of the replacement, refinement and reduction (1,2). As such, the accurate identification of human carcinogens will primarily be based more on in vitro genotoxicity tests without total reliance of in vivo assays to validate results.

Micronuclei (MN) are well-characterised biomarkers of chromosome damage and an established and accepted cytogenetic end point in mutagenicity testing and human population biomonitoring (3–5). MN are found in interphase cells as small, extranuclear bodies resulting from chromosome breaks and whole lagging chromosomes, which are not incorporated into the main nucleus during cell division. MN frequency can be used to screen drug candidates and other test chemicals for both clastogenic and aneugenic potential (3). Indeed, the Organization for the Economical Cooperation and Development (OECD) has published guidelines for the testing of chemicals using the in vitro MN assay (OECD 487) (6).

A novel in vitro human reconstructed skin micronucleus (RSMN) assay that measures MN induced in dividing basal cell keratinocytes of 3D human skin models has recently been developed (7–9). The models’ functional stratum corneum provides more relevant dermal exposure conditions for target cells, avoiding non-physiological concentrations of test chemical that occur in many in vitro tests (8). The RSMN assay may represent a potential replacement for an in vivo genotoxicity assay to support the regulatory safety evaluation for cosmetic products and other dermally exposed chemicals.

The conventional approach for quantitative analysis of micronucleation by visual microscopy is highly labour-intensive, requiring analysis of a large number of cells to obtain statistically reliable data, and can be subjective, leading to unacceptable inter-scorer variability (10,11). Many attempts have thus been made to automate in vitro quantification of MN utilising a variety of methods (12–17). Several systems are available commercially for image analysis of MN, including MetaSystems metafer 18, Imstar Pathfinder Cell Scan (16), GE Healthcare In Cell Analyser (19), Compucyte iCyte automated imaging cytometer (20) and Cellomics Array Scan (21), and for flow cytometric assays, namely the MicroFlow procedure (22). Despite the large numbers of cells that can be rapidly analysed by flow cytometry, there are several shortcomings that limit its applications in the MN assay. For example, identification of MN is based on characteristic distribution on DNA frequency histograms, and as such, other particles including individual chromosomes, chromatin fragments, debris and apoptotic bodies can be erroneously classified as MN (13,22). This may contribute to the large variations in background MN frequency observed in vitro (23). Attempts have been made to
overcome these difficulties however, unlike microscopy-based techniques, no data can be obtained for individual cells using standard flow cytometers, such that individual cells cannot be examined in detail to determine conclusively the presence of MN, and once scored, the samples cannot be reanalysed after long periods have elapsed (24). The main disadvantage of image-based methods is the requirement of some scorer input to minimise MN detection errors, yet gating analysis of cell populations during flow cytometry is also open to subjective evaluation.

This study aimed to optimise, assess and validate the metafer system, which was developed by MetaSystems in 2004, for the automated in vitro cytokinesis-block micronucleus (CBMN) assay in human cell lines commonly used for genotoxicity testing, namely the human lymphoblastoid suspension cell lines (AHH-1, MCL-5 and TK-6) and the human bronchial epithelial adherent cells (BEAS-2B). In this context, a standardised slide preparation protocol was developed to ensure high reproducibility and to minimise false-positive outputs by the system. To validate the optimised scoring procedure, MN frequencies and dose-response curves obtained by conventional visual and automated scoring were compared in AHH-1 cells treated with the direct-acting genotoxins, hydrogen peroxide (H2O2), menadione and potassium bromate (KBrO3). BEAS-2B cells treated with quantum dot nanoparticles and TK-6 cells treated with the DNA crosslinking agent, mitomycin C (MMC) (25). The system was also further adapted for use with the RSMN assay prepared from EpiDerm™ tissues.

Materials and methods

Chemical agents

H2O2, KBrO3, menadione, MMC and cadmium selenide/zinc sulphide, hexadecylamine-coated quantum dots were all purchased from Sigma (Dorset, UK). All chemical dilutions used for treatment of cell line cultures were freshly prepared from stock solutions with water, except quantum dots, which were diluted in dimethyl sulphoxide. Chemical dilutions for treatment of 3D EpiDerm™ skin tissue were freshly prepared in acetone and diluted subsequently in water.

Cell culture

The human lymphoblastoid cell lines, AHH-1 (26) and TK-6, were cultured in RPMI 1640 (Gibco Life Technologies, Paisley, UK) supplemented with 1% t-glutamine (Life Technologies) and 10% donor horse serum (BDGentest, Oxford, UK). MCL-5 human lymphoblastoid B cells were grown as AHH-1, but further supplemented with 200 μg/ml of hygromycin-B (Sigma). BEAS-2B human bronchial epithelial cells were cultured in Dulbecco’s modified Eagle’s medium supplemented with 10% foetal calf serum (Gibco Life Technologies). Cells were maintained at 37°C, in a humidified atmosphere of 5% CO2.

CBMN assay

Chromosomal damage induction was assessed as MN frequency in binucleate (BN) cells. Cell suspensions (10ml) at 1×10^6/ml were seeded for 24h at 37°C, in a 5% CO2 atmosphere. Replicate flasks (n=3, independently produced on different days) were dosed with the appropriately diluted test chemical or negative control (H2O2) for 4h (H2O2, KBrO3 and menadione in AHH-1 cells) or 24h (MMC in TK-6 cells; quantum dots in BEAS-2B cells) after which cells were centrifuged, washed once in phosphate-buffered saline (PBS) and resuspended in 10ml of fresh media containing 4.5 μg/ml (BEAS-2B, MCL-5 and TK-6) or 6 μg/ml (AHH-1) cytochalasin B (Sigma) for one cell cycle (24h BEAS-2B, 22h AHH-1, 20h MCL-5 and 18h TK-6). Cells were harvested into two suspensions and harvested according to visual or metafer procedures. Visual. Treated cells were harvested, cytospun on polished glass slides, fixed in 90% methanol, stained with 20% Giemsa stain solution (20/80ml of phosphate buffer, pH 6.8) and viewed under an Olympus BH2 light microscope with an UplanFl×100 per 1.25 oil objective. Metafer. Treated cells were harvested, resuspended in 10ml of hypotonic solution [0.56% potassium chloride (KCl)] and centrifuged immediately. Cell pellets were resuspended in fixative 1 [methanol:acetic acid:0.9% sodium chloride (5:1:6 parts)] and centrifuged after a 10-min incubation period. Cells were then transferred to fixative 2 [methanol:acetic acid (5:1 parts)] for a 10-min incubation, centrifuged, washed four times and maintained in the final fixative 2-wash at 4°C for 16h. Fixed cells were centrifuged, resuspended in 1ml of fresh fixative 2 and 100 μl of appropriately diluted cell suspension was dropped onto polished, fixed and hydrated slides. Slides were stained with 30 μl of Vectashield mounting medium with 4’-6-diamidino-2-phenylindole (DAPI; Vector Laboratories, Burlingame, USA) and stored in darkness until scanned by the metafer 4 master station.

RSMN assay

Media was aspirated from all wells and replaced with 900 μl of NMM containing 3 μg/ml of cytochalasin B. Tissues were treated with 10 μl of acetone or MMC at 2M on the topical surface. The treatment was repeated 24h later. After the 2x24h treatment period, tissues were removed from media and submerged in 5ml of Dulbecco’s phosphate-buffered saline (DBPS; Life Technologies) for 15 min, followed by 5ml of 1g/l EDTA (diluted in DBPS from 0.5M EDTA; Ambion Life Technologies) for 15 min. Tissues were then exposed to trypsin (Life Technologies) for 15 min. The tissue was then carefully lifted from the supporting membrane using fine forceps and trypsinised further to individualise basal cells. The tissue and tissue cup were then discarded and the resulting cell suspension was added to 1ml of trypsin-neutralising solution (Dulbecco’s modified Eagle’s medium; Gibco Life Technologies) supplemented with 10% foetal bovine serum (Gibco Life Technologies).

Samples were centrifuged, supernatant was discarded and pellet was resuspended with gentle agitation, before 1ml of pre-warmed 0.075M KCl was added slowly. After 3min in KCl, 3ml pre-cooled (4°C) methanol/acetic acid (3:1) was added, and samples were mixed gently prior to centrifugation. Supernatant was removed immediately, the pellet resuspended and divided into two suspensions. Slides were prepared for either visual or automated scoring procedures. Visual. Cell suspension was gently pipetted in a concentrated area on a slightly angled, dry, clean glass slide. Slides were left to air-dry prior to staining in 40 μg/ml acridine orange (Sigma-Aldrich) and diluted in DBPS, for 3min. Slides were then washed in 0.05M staining buffer (pH 6.8; Intravention) for 30min, air-dried and viewed using a fluorescent microscope equipped with a blue filter.

Metafer. Samples were stored overnight at 4°C in KCl/fixative (methylene/ acetic acid). Following storage, samples were centrifuged, supernatant was removed, pellets resuspended in fresh fixative, and a 100 μl of the pellet suspension dropped onto polished, fixed and hydrated slides. Slides were stained and scored as detailed in CBMN assay metafer section.

Scanning

Slides were scanned at x10 magnification with the metafer 4 master station, comprising a Carl Zeiss Axioplan Imager Z1, equipped with a Maerzhaeuser stepping motor stage that scans eight slides unattended. The metafer classifi- fier (Table I) was run on the metafer MSsearch platform version 3.5 software (MetaSystems, Altusseheim Germany). Images were acquired on a high-resolution megapixel charge-coupled device (CCD) camera (Carl Zeiss). A predetermined scan area was set to accommodate the length of the slide and to avoid the outside margins of the cell preparation area. The plane of focus was determined at a number of grid positions distributed evenly across the scan area. The system requires the selection of a reference object within a small centre cross-haired square, which will be utilised as a slide position reference and as a reference field for fluorescent signal detection. To enable a correct integration time of the CCD camera and correct signal level measurement, the reference area should contain objects with typical staining and not fields containing artefacts with very high staining intensity. The number of BN cells required was pre-set on the system, such that when this number of cells was achieved by the scan, the system would move on to the next slide without scanning the remaining area of the slide. The classifer used for the detection of BN and MN was developed in association with MetaSystems and Zeiss.
**Table I. Metafer classifier parameters**

<table>
<thead>
<tr>
<th></th>
<th>Old classifier</th>
<th>New optimised classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclei parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Image processing operations</td>
<td>Sharpen (3,4)</td>
<td>Sharpen (3,2) MedianV (3) MedianH (3)</td>
</tr>
<tr>
<td>Object threshold (%)</td>
<td>40</td>
<td>30</td>
</tr>
<tr>
<td>Minimum area (µm²)</td>
<td>10.0</td>
<td>20.0</td>
</tr>
<tr>
<td>Maximum area (µm²)</td>
<td>400.00</td>
<td>400.0</td>
</tr>
<tr>
<td>Maximum relative concavity depth</td>
<td>0.900</td>
<td>0.900</td>
</tr>
<tr>
<td>Maximum aspect ratio</td>
<td>1.500</td>
<td>1.500</td>
</tr>
<tr>
<td>Maximum distance (µm)</td>
<td>30.0</td>
<td>30.0</td>
</tr>
<tr>
<td>Maximum area asymmetry (%)</td>
<td>70</td>
<td>70</td>
</tr>
<tr>
<td>Region of interest radius (µm²)</td>
<td>40.0</td>
<td>40.0</td>
</tr>
<tr>
<td>Maximum object area in ROI (µm²)</td>
<td>50.0</td>
<td>50.0</td>
</tr>
<tr>
<td>MN parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Image processing operations</td>
<td>MedianV (3) MedianH (3)</td>
<td>MedianV (3) MedianH (3)</td>
</tr>
<tr>
<td>Object threshold (%)</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td>Minimum area (µm²)</td>
<td>1.00</td>
<td>1.50</td>
</tr>
<tr>
<td>Maximum area (µm²)</td>
<td>55.00</td>
<td>55.0</td>
</tr>
<tr>
<td>Maximum relative concavity depth</td>
<td>1.000</td>
<td>0.900</td>
</tr>
<tr>
<td>Maximum aspect ratio</td>
<td>3.500</td>
<td>4.000</td>
</tr>
<tr>
<td>Maximum distance (µm)</td>
<td>30.0</td>
<td>25.0</td>
</tr>
<tr>
<td>Autoseparate</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>Metafer 4 software outputs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Artefactual-positive MN</td>
<td>4.17±3.73%</td>
<td>2.62±2.58%</td>
</tr>
<tr>
<td>Artefactual-negative MN</td>
<td>0.37±0.31%</td>
<td>0.34±0.22%</td>
</tr>
</tbody>
</table>

The original classifier presented by MetaSystems and the optimised classifier for the detection of BN cells and BN cells containing MN using a ×10 objective are shown.

**Metafer classifier**

The classifier for scanning BN preparations was set to the criteria detailed in Table I for scoring utilizing a ×10 objective. The same classifier was utilised for CBMN analysis of all of the different cell lines tested (AHH-1, BEAS-2B, MCL-5 and TK-6). Parameters described in the nuclei group define how the system selects the BN cells to be analysed. The parameters set in the MN group define how the system detects the MN in the selected BN cells. The main changes to the classifier involved altering the sequence of image processing operations applied after image capturing. For example, MN background image processing was updated to include the operation SBHistoMax (subtract background histogram maximum), where the grey level histogram is computed, the histogram maximum is detected and the position of the maximum is then subtracted from every image grey level, resulting in cleaner image backgrounds in the final output. The minimum area in square micrometers for a single nucleus and MN to be accepted for analysis was also optimised.

**Scoring procedure**

The criteria for identifying MN were as previously described by Fenech (27). A minimum of 2000 BN cells were scored per replicate slide prepared from each cell line, and each dose was performed in triplicate (an average of 6000 BNs per dose, except for slides prepared for visual analysis of KBrO₃, mene-dione, MMC and quantum dots, which were performed in duplicate (an average of 40 000 BN per dose)). A 1000 BN cells were scored per replicate slide prepared from EpiDerm™ basal cells, and each treatment was performed in triplicate (an average of 3000 BN per treatment). During metafer CBMN assay scoring, all cells captured by the system can be viewed in an image gallery. The image gallery was ordered by entering the general gallery set up and electing the microscope’s ×100 objective, and MN-negative cells were scanned based on the gallery images on the screen. Cells that appeared as false negatives were checked at ×100 magnification. The error rate of identification of true BN cells (i.e. two nuclei in a BN cell should have intact nuclear membranes, should be approximately equal in size, may touch but ideally should not overlap each other, etc.) by the metafer system was <1%.

**Cytoxicity and cytostasis**

Satellite cultures were seeded as previously described for the CBMN assay. Cells were counted on the Coulter counter (Beckman) 1 h before dosing with appropriate test chemical (starting count). Cells were then centrifuged, washed once in PBS and resuspended in 10 ml of fresh media without cytochalasin B for one cell cycle (22 h AHH-1 and BEAS-2B, 20 h MCL-5 and 18 h TK-6). Cell counter readings were performed for a second time (final count) and the relative population doubling (RPD) value was calculated using the following formula:

\[
\text{RPD} = \frac{\text{Number of population doublings in treated cultures}}{\text{Number of population doublings in control cultures}} \times 100, \quad \text{where population doubling} = \left[ \log \left( \frac{\text{post-treatment cell number}}{\text{initial cell number}} \right) \right]/0.34 \text{ or relative increase in cell counts (RICC):}
\]

\[
\text{RICC} = \frac{\text{Increase in number of cells in treated cultures (final – starting)}}{\text{Increase in number of cells in control cultures (final – starting)}} \times 100.
\]

All cell counts were taken from the average of two particle counter readings.

**Optimisation of metafer slide preparation**

To optimise the optimal amount of DAPI stain required for the metafer protocol, control slides were prepared using AHH-1, BEAS-2B, MCL-5 and TK-6 cells as previously described for the CBMN metafer assay. Prepared slides (see CBMN metafer assay) were stained with varying amounts of DAPI vectashield solution (30 or 75 µl). The stain was applied at 3 points along the length of the slide (either 3x10 or 3x25 µl) before applying a 300x600 mm coverslip. Staining was performed in triplicate, with a total of 3000 BN scored for each experiment. To ascertain the optimal slide cell density for the metafer protocol, fixed TK-6 cell suspensions were prepared from control flasks as described for the CBMN metafer assay, but at a concentration of 3x10⁶ cells/ml. Serial dilutions of this cell suspension were prepared at the following densities: 1x10⁶, 1x10⁷, 8x10⁶, 6x10⁶, 5x10⁶, 4x10⁶, 3x10⁶ cells/ml, and 100 µl of diluted cell suspension was dropped onto fixed, hydrated and polished slides. The procedure of dropping cells onto slides enables an even distribution of cells across
the length of the slide. Clumping of cells is prevented during the harvesting procedure by ensuring the cell pellets are fully resuspended before adding fixative. The experiment was performed in triplicate, with a minimum of 6000 BN scored per cell density measure.

Statistical analysis
Dose-response curves were analysed using a one-way analysis of variance, followed by a Dunnett’s post hoc test, to determine if any of the treatment doses were significantly different from the zero dose. When CBMN assay experiments were performed in duplicate, Fisher’s exact test was performed to determine any differences in dose from background levels. Fisher’s exact test was utilised to ascertain any differences in experimental optimisation protocols. Dose-response curves of the H2O2, KBrO3, menadione and MMC were analysed mathematically utilising the benchmark dose (BMD) approach. This modelling approach differs from the no observed genotoxic effect level (NOGEL) in that it provides an estimate of the size of the effect associated with the estimated dose and estimates a dose (i.e. the BMD) that produces a predetermined increase in the response over the controls (28). The lower limit of the one-sided 95% confidence interval on the BMD is termed the BMDL. A BMDL05 refers to the estimate of lower 95% confidence interval of a dose that produces a 10% increase over the fitted background level for continuous end points. BMDL05 values were derived using the dose-response modelling software package PROAST, developed at the National Institute for Public Health and the Environment (RIVM) in the Netherlands (www.proast.nl; versions 26.4, 28.1 and 28.3) (29). Comparisons of manual versus automated MN dose-response results were performed using linear regression analysis.

Results
Initial automated analysis of untreated, control slides, demonstrated that the metafer 4 classifier required further adaptation due to the occurrence of several shortfalls including the resolution of detection of small MN and MN situated in close proximity to nuclei. The aim of this optimization process being to develop an automated platform for the detection and analysis of BN cells and MN-containing BN cells prepared from a variety of different suspension and adherent cell lines by the metafer system. By altering the classifier parameters (with the guidance of MetaSystems), most of these difficulties were overcome (see Table 1), such that levels of artefactual positives scored by the system reduced from 4.17±3.73% to 2.62±2.58% and levels of artefactual-negative MN missed by the system decreased from 0.37±0.31% to 0.34±0.22% (data not shown). To develop an optimal automated protocol for the CBMN assay utilising the system, firstly the slide preparation was optimised in relation to staining intensity and cell density, as these factors greatly influence the output of the software and the efficiency of the scoring system. This protocol was then validated by comparing the dose-response curves and MN frequencies of the genotoxic chemicals H2O2, KBrO3 and menadione in AHH-1 cells, quantum dots in BEAS-2B cells and MMC in TK-6 cells of DAPI-stained slides primed for metafer automated analysis versus Giemsa-stained slides prepared for conventional visual light microscope analysis. Additionally, the ability of the metafer protocols and classifier to perform the RSMN assay utilising slides prepped from 3D EpiDerm™ human skin models was assessed by comparing data from the metafer to results from visual scoring in cells harvested from the same individual tissues.

Optimisation of metafer slide preparation
To determine the optimum slide DAPI stain intensity, the quantity of stain solution applied to the slides was altered and the metafer output recorded. Slides stained with the lesser staining intensity (30 µl) of DAPI had considerably less artefactual positives reported by the metafer software than slides with the higher staining intensity (75 µl) for each of the four cell lines: AHH-1, BEAS-2B, MCL-5 and TK-6 (Figure 1). Improvements of 72–95% were recorded for the elimination of artefactual-positive outputs when less DAPI staining was applied to slides and this difference was deemed statistically significant by Fisher’s exact test (P = 0.00) for each of the cell lines analysed.

Altering slide cell density by serially diluting a cell suspension of known cell quantity influenced output by the metafer software; the highest cell density increased artefactual-positively output to 13% from 2.8%, which was observed for the lowest cell density, and this difference was deemed significant by Fisher’s exact test (P = 0.001). Altering cell density also affected scorer output, with a lower density resulting in a reduced level of reported MN by the scorer (Figure 2). For example, the mean frequency of MN observed on control slides decreased significantly from 1.3±0.18% when slides were prepared from cell suspensions of 1.2×106 cells/ml to 0.7±0.05% on slides prepared from cell suspensions at a density of 3×106 cells/ml (P = 0.00). Higher levels of chromosomal damage observed on slides prepared with the cell suspensions containing the highest number of cells could be associated with a greater scorer misinterpretation due to an increased difficulty to resolve MN when cells are too crowded upon slides. Examples of hard to resolve events present on highly dense slides are shown in Figure 3. The optimal density and dispersion of cells for metafer examination were, therefore, achieved.

Fig. 1. The importance of optimal nuclear staining on false-positive MN detection. Control slides were prepared for the CBMN assay for the human cell lines: AHH-1, BEAS-2B, MCL-5 and TK-6. The slides were treated with either 30 or 75 µl of DAPI stain and the artefactual-positive output of the metafer 4 system recorded. %Mn/Bn, percentage of micronucleated BNs; columns, average artefactual-positive output from triplicate slides; bars, standard deviation; asterisk, statistically significant increase in false-positive MN/BN, P = 0.00.
Evaluation of metafer micronucleus assay protocol

Several minor difficulties were encountered during the automated scoring process of the CBMN assay, mainly due to the loss of cytoplasmic boundaries during slide preparation. Inspection of all BN cell images during the lengthy optimisation process revealed levels of 2.62±2.58% artefactual-positive MN scored automatically by the system and levels of 0.34±0.22% artefactual-negative MN detected (data not shown). Artefactual micronucleated cells (MN/BN) may be reported by the system due to debris, chromosomes, trinuclear cells containing one smaller nuclear body and apoptotic cells positioned in close proximity to the BN cells. Some examples of artefactual-positive outputs by the metafer system are depicted in Figure 5. The HUMN scoring criteria were followed to identify MN; however, some allowances were needed (30). For example, close attention to the staining patterns of the nuclei was necessary to prevent the identification of two mononuclear cells as a BN cell. The distance of MN from the nuclei must also be taken into consideration. A semi-automated method of scoring was found most successful for minimising artefactual-positive and artefactual-negative MN identification. Firstly, cellular images captured by the system were ordered in the gallery by descending MN frequency such that aberrant cells (MN positive) were present at the top of the gallery irrespective of where they were positioned on the slide. MN-positive cells were then confirmed visually by the scorer using ×100 magnification and MN-negative images from slides prepared from cell suspensions of 3–6×10⁵ cells/ml. Examples of output slide density images by the metafer software are depicted in Figure 4.

**Validation of metafer protocol**

Several minor difficulties were encountered during the automated scoring process of the CBMN assay, mainly due to the loss of cytoplasmic boundaries during slide preparation. Inspection of all BN cell images during the lengthy optimisation process revealed levels of 2.62±2.58% artefactual-positive MN scored automatically by the system and levels of 0.34±0.22% artefactual-negative MN detected (data not shown). Artefactual micronucleated cells (MN/BN) may be reported by the system due to debris, chromosomes, trinuclear cells containing one smaller nuclear body and apoptotic cells positioned in close proximity to the BN cells. Some examples of artefactual-positive outputs by the metafer system are depicted in Figure 5. The HUMN scoring criteria were followed to identify MN; however, some allowances were needed (30). For example, close attention to the staining patterns of the nuclei was necessary to prevent the identification of two mononuclear cells as a BN cell. The distance of MN from the nuclei must also be taken into consideration. A semi-automated method of scoring was found most successful for minimising artefactual-positive and artefactual-negative MN identification. Firstly, cellular images captured by the system were ordered in the gallery by descending MN frequency such that aberrant cells (MN positive) were present at the top of the gallery irrespective of where they were positioned on the slide. MN-positive cells were then confirmed visually by the scorer using ×100 magnification and MN-negative images from slides prepared from cell suspensions of 3–6×10⁵ cells/ml. Examples of output slide density images by the metafer software are depicted in Figure 4.

**Fig. 2.** The importance of optimal cell density for MN detection. Effect of increasing cell density on scorer identification of MN on TK-6 control slides. %Mn/Bn, percentage of micronucleated BNs; columns, average of three slides; bars, standard deviation; asterisk, statistically significant increase in MN frequency from the lowest cell density, $P = 0.001$.

**Fig. 3.** Metafer 4 gallery images showing examples of hard to resolve events. Images captured on metafer 4 system from DAPI-stained slides prepared from TK-6 cells showing examples of hard to resolve events. At high cell densities, resolution of MN becomes increasingly difficult due to the lack of cytoplasmic boundaries.
scanned using the captured images present on the screen. As previously discussed, slide preparation is paramount and it is important to ensure an even distribution of cells across the length of the slide at the correct density. Poor slide preparation, with clumping of cells, or cells at high density can interfere with the correct identification of MN/BN by the system and by the scorer. Despite the requirement for user input to correctly identify MN, the metafer system represents a high content, high throughput, time-saving alternative to conventional scoring techniques, cutting scoring time by two-thirds from ~45 min (visual scoring) to 15 min (automated scoring, which includes the systems’ scanning).

To validate the optimised semi-automated methodology, comparisons were made between the metafer system and conventional light microscopy, after exposure to a variety of direct-acting DNA damaging agents, utilising some automated data previously published by Seager et al. (25). Comparison of the frequencies of MN/BN for the four compounds tested showed that both scoring methods detect a similar dose response for the chemicals analysed (Figure 6). Cytotoxicity was at acceptable levels (i.e. <55 ±5%, as recommended by OECD 487 guideline) for pro-oxidant chemicals (H₂O₂, KBrO₃, and menadione) and at lower doses of MMC (Figure 6). Lower frequencies of MN were reported by conventional visual scoring as compared to the automated system, e.g. average background levels of MN/BN in untreated cultures were 1.39 ± 0.37% and 0.62 ± 0.31% for H₂O₂, 1.23 ± 0.25% and 0.5 ± 0.0% for KBrO₃, 1.16 ± 0.04% and 0.65 ± 0.14% for menadione, 1.38 ± 0.08% and 0.54 ± 0.24% for MMC and 1.63 ± 0.15% and 1.0 ± 0.1% for quantum dots, for automated and visual scoring, respectively. This may reflect the increased sensitivity of the metafer system to detect MN. Despite this, concentrations that produced the first significant increase in MN frequency above the background level were generally the same for both methods, and shapes of the dose-response curves mirrored each other extremely well. Interestingly, where there are differences in the lowest observed genotoxic effect levels (LOGELs) of visual versus automated, the metafer tends to be more sensitive generating lower LOGELs. Statistical analysis of dose responses produced via the automated and visual methods gave LOGELs of 25 µM for H₂O₂ (P = 0.03 and P = 0.000 for automated and visual scoring, respectively), 0.6 mM (P = 0.007) by automated methods and 0.5 mM (P = 0.024) by visual methods for KBrO₃, 2.9 µM (P = 0.036) by automated methods and 5.8 µM (P = 0.0006) by visual scoring menadione, and 0.67 mM by automated scoring and 2.67 mM by visual methods for MMC (P = 0.0007 and P = 0.031 for automated and visual scoring, respectively) (Figure 6). Treatment with quantum dots showed no significant increase in dose response and thus no LOGEL is reported. Mathematical modelling of each of the separate dose responses utilising the BMD approach produced the BMD₁₀ and lower confidence intervals (BMDL₁₀) shown in Table II. Both reported outputs (BMDL₁₀ and BMD₁₀) show tight correlation between visual and automated scoring for each chemical.

Further, linear regression analysis comparing the similarity of the dose-response curves and MN frequencies generated by the two different methods reported highly positive correlations between the shape of the automated and manual dose-response outputs with R² = 0.70 for H₂O₂, R² = 0.89 for KBrO₃, R² = 0.98 for menadione, R² = 0.99 for MMC and R² = 0.68 for quantum dots. The regression analysis performed found the calculated F to be higher than the critical F when comparing the two dose-response curves for each chemical, providing significant statistical evidence for a linear relationship between the two scoring techniques P = 0.001 or less. Thus, dose responses were shown to be comparative between visual and automated, and as such, both methodologies are equally suitable for dose-response analysis in the low dose region.

To assess whether MN frequency could be scored effectively using the metafer in the 3D EpiDerm™ human skin models, results from the metafer were compared to results from visual scoring in cells harvested from the same individual tissues (Figure 7). No statistically significant differences were detected between visual and metafer scoring methodologies for the same test article by t test, P = 0.12 and P = 0.42 for acetone and MMC, respectively. However, the metafer appeared to slightly
over-estimate MN frequency relative to visual for three replicates, consistent with data observed for the dose responses of the chemicals analysed.

Discussion

Although the MN assay is an efficient, relatively simple and well-validated method, visual scoring of MN is a laborious and time-consuming task, which requires subjective analysis of large numbers of cells and donors to obtain statistically relevant data (14). There has been strong interest in automating the measurement of the MN end point over the past decade in order to reduce labour and costs, and enhance throughput, scoring objectivity and statistical power of the assay particularly for dose-response analysis. We aimed to optimise and implement an automated CBMN assay facility using fluorescence-microscopy-based detection and a commercially available program for metaphase recognition developed by MetaSystems, for use in a range of human cell lines. The metafer image analysis system was able to identify damaged cells from slides prepared from both suspension cell lines (AHH-1, MCL-5 and TK-6) and adherent cultures (BEAS-2B), as well as slides from the RSMN assay. Manual assessment of the captured images enabled MN to be quickly distinguished from other cellular debris. Once the slide preparation and staining were optimised to minimise artefactual-positive outputs, metafer allowed comparable scoring of MN to traditional, visual methods, providing highly correlated trends in observed DNA damage induction for all five of the test chemicals analysed. Higher background levels of MN/BN were observed for the metafer system versus visual scoring, and this may reflect the increased sensitivity of the semi-automated assay, yet BMD values were tightly correlated between the two scoring techniques suggestive that the metafer may be employed successfully in scoring low dose regions without significant over- or under-prediction of chromosomal damage relative to visual scores.

Although the metafer system can be utilised as a fully automated platform, the protocol used in this study was semi-automated with a two-step procedure. First, automatic detection of MN-positive BN cells was performed by the metafer 4 program, followed by manual substantiation on screen by the scorer. At present, it is essential to inspect all gallery pictures to remove artefactual-positive MN that are scored automatically and identify false-negative cells, which are occasionally missed by the system. The automatic errors arise through the inability to identify the cytoplasm of damaged cells (the assignment of nuclei and MN to a cell relies on the use of other parameters and algorithms such as similarity of nuclei and distance between nuclei), as well as identification of background signals. Although the interactive use of the program in this way requires some operator time, the automated test remains much faster than conventional scoring techniques, taking approximately a third of the time to score a single slide (~15 min) and represents a technological advance over conventional, visual scoring methods.

The main advantages of the metafer system over flow-cytometry-based methods are the potential to distinguish between MN by using a combination of DAPI staining with other fluorescent probes (e.g. centromeric or telomeric). This would provide the opportunity to obtain additional information about the mechanism of MN formation and mechanism of action of the test chemical (aneugenic versus clastogenic). In this respect, the ability of the system to view MN-positive cells rapidly would prove extremely advantageous. Other groups using the metafer have also utilised the system to identify apoptotic, necrotic and fragmented cells and provide details of cellular cytotoxic damage responses, as MN may be distinguished readily from cellular debris (31).

Fig. 5. Metafer 4 captured images of micronucleated cells (MN/BN). (A) Image of a positive MN/BN. (B–D) Examples of artefactual-positive outputs by the system, where (B) the distance of the MN from nuclei is too great, and (C and D) staining intensity and distance between nuclei indicate these cells are not BNs.
Table II. NOGEL, BMD and BMDL_{10} values for MN end points induced by H_{2}O_{2}, KBrO_{3}, menadione and MMCF

<table>
<thead>
<tr>
<th>Treatment</th>
<th>NOGEL</th>
<th>BMD_{10}</th>
<th>Ratio</th>
<th>BMDL_{10}</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>H_{2}O_{2} (µM)</td>
<td>Visual</td>
<td>20</td>
<td>7.36</td>
<td>1.87</td>
<td>3.29</td>
</tr>
<tr>
<td></td>
<td>Metafer</td>
<td>20</td>
<td>7.36</td>
<td>1.87</td>
<td>3.29</td>
</tr>
<tr>
<td>KBrO_{3} (mM)</td>
<td>Visual</td>
<td>0.4</td>
<td>1.12</td>
<td>0.67</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>Metafer</td>
<td>0.5</td>
<td>1.12</td>
<td>0.67</td>
<td>0.65</td>
</tr>
<tr>
<td>Menadione (µM)</td>
<td>Visual</td>
<td>1.45</td>
<td>0.75</td>
<td>0.67</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>Metafer</td>
<td>2.9</td>
<td>1.12</td>
<td>0.67</td>
<td>0.65</td>
</tr>
<tr>
<td>MMC (nM)</td>
<td>Visual</td>
<td>0</td>
<td>3.32</td>
<td>2.31</td>
<td>1.65</td>
</tr>
<tr>
<td></td>
<td>Metafer</td>
<td>2.0</td>
<td>1.44</td>
<td>2.31</td>
<td>1.65</td>
</tr>
</tbody>
</table>

Ratio shown as italic values indicate fold difference between automated (metafer) and visual BMD_{10} and BMDL_{10} metrics.
Background levels of MN were reported on average as 1.32±0.26% between all of the cell lines and chemicals analysed, highlighting the high reproducibility and low variation of this technique. In addition, all images captured are stored by the system with their specific slide coordinates, which allows the user to revisit cells of interest on the slide, based on the image bank. Thus, slides can be reloaded if required, and if necessary, repeated scoring and peer review of the same slide are possible. In contrast, with the automated flow cytometry, MN assay measured samples cannot be stored for retrospective studies (e.g. for analysis with other probes), confirmation of correct scorer analysis or archival preservation (13). Moreover, the system and the classifier settings are easily adaptable to different cell types; the metafer classifier employed in this study has been applied to several different cell lines, both suspension and adherent cells, whilst other groups have reported the use of the system with other cell types and lines (15,18,31).

We have also demonstrated that the metafer may be employed successfully in scoring MN frequency in cells harvested from the 3D EpiDermtm human skin models using the RSMN assay. This RSMN assay is a promising new in vitro genotoxicity test designed to evaluate chromosomal damage induced by dermally applied chemicals (7). Following recent bans for testing of cosmetic ingredients in animals, it provides a potential alternative to in vivo-based genotoxicity tests (8). A particularly valuable finding is that the shortened harvesting protocol recommended for the RSMN assay is compatible with both DAPI staining and the metafer. This suggests that the multiple wash steps in the currently recommended harvesting protocol for the metafer may be superfluous, and use of the shortened protocol provides greater convenience without compromising results. This allows a more rapid fixation procedure without wash steps, which may also prevent loss of cells during the harvesting procedure. Further analysis utilising this shortened protocol with other cell lines are required.

Conclusion

In conclusion, the semi-automated metafer system provides a reliable, accurate and efficient method for the scoring of the in vitro CBMN assay in a range of human cell lines. The development of a standardised protocol and classifiers, for use within different laboratories for the assessment of MN by automated image analysis, would be advantageous to determine inter-laboratory variation rates, with the aim of regulating the whole procedure.

Funding


Acknowledgements

The authors would like to thank Mrs Sally James for all the time and effort contributed to aiding in the development of the optimised protocol.

References

18. Varga, D., Johannes, T., Jainta, S., Schuster, S., Schwarz-Boeger, U.,
procedure for the micronucleus test by image analysis. Mutagenesis, 19,
391–397.
Evaluation of high-throughput screening for in vitro micronucleus test
using fluorescence-based cell imaging. Mutagenesis, 26, 709–719.
(1999) Laser-scanning cytometry: a new instrumentation with many appli-
Evaluation of an automated in vitro micronucleus assay in CHO-K1 cells.
in cell cultures and human lymphocytes: advantages and disadvantages.
23. Bryce, S. M., Avlasevich, S. L., Bemis, J. C., Phonethepswath, S. and
assay represents an efficient tool for comprehensively characterizing geno-
In vitro micronucleus scoring by flow cytometry: differential staining of
micronuclei versus apoptotic and necrotic chromatin enhances assay reli-
DNA damage in human lymphoblastoid cells: homeostatic mechanisms of
lymphoblast line, AHH-1, competent for xenobiotic metabolism. Mutat.
Protoc., 2, 1084–1104.
approaches for assessing dose-response relationships in genetic toxicology
30. Fenech, M., Chang, W. P., Kirsch-Volders, M., Holland, N., Bonassi, S.
and Zeiger, E. (2003) HUMN project: detailed description of the scoring
criteria for the cytokinesis-block micronucleus assay using isolated human
A rapid, semi-automated method for scoring micronuclei in mononucleated