# An increased micronucleus frequency in peripheral blood lymphocytes predicts the risk of cancer in humans

Stefano Bonassi[1,*], Ariana Znaor[3], Marcello Ceppi[1], Cecilia Lando[1], Wushou Peter Chang[4], Nina Holland[5], Micheline Kirsch-Volders[6], Errol Zeiger[7], Sadayuki Ban[8,22], Roberto Barale[9], Maria Paola Bigatti[10], Claudia Bolognesi[2], Antonina Cebulska-Wasilewska[11], Eleonora Fabianova[12], Alexandra Fucic[13], Lars Hagmar[14,†], Gordana Joksic[15], Antonietta Martelli[16], Lucia Migliore[9], Ekaterina Mirkova[17], Maria Rosaria Scarfi[18], Andrea Zijno[19], Hannu Norppa[20] and Michael Fenech[21]

[1]Unit of Molecular Epidemiology, [2]Unit of Environmental Carcinogenesis, National Cancer Research Institute, Genoa, Italy, [3]Croatian National Cancer Registry, Croatian National Institute of Public Health, Zagreb, Croatia, [4]Institute of Environmental Health Sciences, National Yang Ming University Medical School, Taipei, Taiwan, [5]School of Public Health, University of California, Berkeley, CA, USA, [6]Laboratory for Cell Genetics, Vrije Universiteit Brussel, Brussel, Belgium, [7]Errol Zeiger Consulting, Chapel Hill, NC, USA, [8]Department of Radiobiology/Molecular Epidemiology, Radiation Effects Research Foundation, Hiroshima, Japan, [9]Dipartimento di Scienze dell'Uomo e dell'Ambiente, University of Pisa, Pisa, Italy, [10]Dipartimento di Biologia Animale, University of Turin, Turin, Italy, [11]Chair of Epidemiology and Preventive Medicine, Jagiellonian University Medical College, Kraków, Poland, [12]State Health Institute, Banská Bystrica, Slovakia, [13]Institute for Medical Research and Occupational Health, Zagreb, Croatia, [14]Department of Occupational and Environmental Medicine, Lund University, Lund, Sweden, [15]Vinca Institute of Nuclear Sciences, Medical Protection Centre, Belgrade, Yugoslavia, [16]DIMI, University of Genoa, Genoa, Italy, [17]National Center of Public Health Protection, Sofia, Bulgaria, [18]CNR-IREA, Naples, Italy, [19]Department of Environment and Primary Prevention, Istituto Superiore di Sanità, Rome, Italy, [20]New Technologies and Risks, Finnish Institute of Occupational Health, Helsinki, Finland, [21]CSIRO Human Nutrition, Adelaide, Australia and [22]Present address: National Institute of Radiological Sciences, Chiba, Japan

*To whom correspondence should be addressed at: Unit of Molecular Epidemiology, National Cancer Research Institute, Largo R. Benzi, 10, 16132-I, Genoa, Italy. Tel: +390 10 5600 924; Fax: +390 10 5600 501; Email: stefano.bonassi@istge.it

The frequency of micronuclei (MN) in peripheral blood lymphocytes (PBL) is extensively used as a biomarker of chromosomal damage and genome stability in human populations. Much theoretical evidence has been accumulated supporting the causal role of MN induction in cancer development, although prospective cohort studies are needed to validate MN as a cancer risk biomarker. A total of 6718 subjects from of 10 countries, screened in 20 laboratories for MN frequency between 1980 and 2002 in *ad hoc* studies or routine cytogenetic surveillance, were selected from the database of the HUman MicroNucleus (HUMN) international collaborative project and followed up for cancer incidence or mortality. To standardize for the inter-laboratory variability subjects were classified according to the percentiles of MN distribution within each laboratory as low, medium or high frequency. A significant increase of all cancers incidence was found for subjects in the groups with medium (RR = 1.84; 95% CI: 1.28–2.66) and high MN frequency (RR = 1.53; 1.04–2.25). The same groups also showed a decreased cancer-free survival, i.e. $P = 0.001$ and $P = 0.025$, respectively. This association was present in all national cohorts and for all major cancer sites, especially urogenital (RR = 2.80; 1.17–6.73) and gastro-intestinal cancers (RR = 1.74; 1.01–4.71). The results from the present study provide preliminary evidence that MN frequency in PBL is a predictive biomarker of cancer risk within a population of healthy subjects. The current wide-spread use of the MN assay provides a valuable opportunity to apply this assay in the planning and validation of cancer surveillance and prevention programs.

## Introduction

Measurement of micronucleus (MN) frequency in peripheral blood lymphocytes (PBL) is extensively used in molecular epidemiology and cytogenetics to evaluate the presence and the extent of chromosomal damage in human populations exposed to genotoxic agents or bearing a susceptible genetic profile (1). This assay has been also successfully applied to identify dietary and genetic factors that have a significant impact on genome stability (2). The high reliability and low cost of the MN technique, has contributed to the worldwide success and adoption of this biomarker for *in vitro* and *in vivo* studies of genome damage (3).

MN originate from chromosome fragments or whole chromosomes that are not included in the main daughter nuclei during nuclear division (Figure 1). The formation of MN in dividing cells is the result of chromosome breakage due to unrepaired or mis-repaired DNA lesions, or chromosome malsegregation due to mitotic malfunction. These events may be induced by oxidative stress, exposure to clastogens or aneugens, genetic defects in cell cycle checkpoint and/or DNA repair genes, as well as deficiencies in nutrients required as co-factors in DNA metabolism and chromosome segregation machinery (2,4–8). All these events can cause the formation of MN through chromosomal rearrangements, altered gene expression or aneuploidy, effects associated
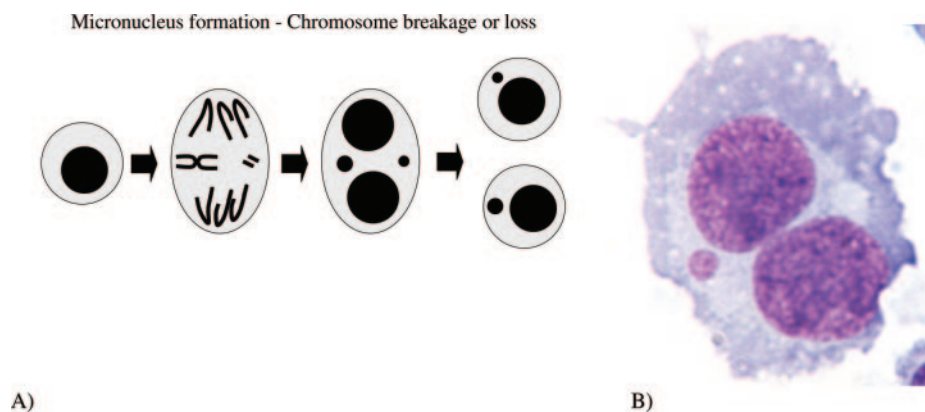
**Fig. 1.** (**A**) Schematic diagram showing the origin of MN from either a lagging chromosome fragment or a whole chromosome. (**B**) A photomicrograph of a mitogen-stimulated, cytokinesis-blocked lymphocyte containing one MN.

with the chromosome instability phenotype often seen in cancer (5,9–11).

The presence of an association between MN induction and cancer development is supported by a number of observations. The most substantiated include: (i) the high frequency of this biomarker in untreated cancer patients and in subjects affected by cancer-prone congenital diseases, e.g. Bloom syndrome or ataxia telangiectasia (1,9); (ii) the presence of elevated MN frequencies in oral mucosa, used as a surrogate biomarker of cancer in clinical chemoprevention trials (12); (iii) the correlation existing between genotoxic MN-inducing agents and carcinogenicity, e.g. ionizing and ultraviolet radiation (13,14); and (iv) the inverse correlation between MN frequency and the blood concentration and/or dietary intake of certain micronutrients associated with reduced cancer risk, such as folate, calcium, vitamin E and nicotinic acid (8). Further evidence—based on the mechanistic and experimental correlation existing between chromosomal aberrations (CA) and MN (1,9)—comes from the results of recent cohort studies, which in most cases demonstrated that the frequency of CA in PBL of healthy subjects is a predictor of cancer risk (15–19). The possible association of lymphocyte MN frequency with cancer risk has earlier been examined in Swedish and Italian cohorts (15,16), although no conclusions could be drawn from those studies because of the young age of the cohorts and the small number of events (18 cancer cases and 9 cancer deaths).

To test the hypothesis that PBL MN are predictive of increased cancer risk, we assembled a large international cohort of subjects whose lymphocytes had been screened for MN frequency between 1980 and 2002 and who were free of cancer at the time of testing. The study was conducted within the framework of the HUman MicroNucleus project (HUMN), an international collaborative project, which allowed gathering of data on 6718 individuals studied in 20 cytogenetics laboratories from 10 countries (1).

## Subjects and methods

The international collaborative HUMN project was developed to improve knowledge of the biology and relevance of MN induction and its application to human population studies (1). Details about this initiative and the list of publications produced under the HUMN project can be found on the project website, www.HUMN.org. The study was also part of the *Cytogenetic Biomarkers and Human Cancer Risk* project (CancerRiskBiomarkers), a European collaborative research project funded by the European Union.

About 50 laboratories actively involved in the HUMN project or included in the mailing list were invited to contribute their databases to the cohort study, assuming that personal identification was available for all subjects and a link with local or national cancer registry was possible. An arbitrarily chosen minimum size of 100 subjects analysed in the same laboratory (even in different studies) was required to be eligibile for inclusion in the cohort. A detailed description of the protocol used for measurement of MN frequency in PBL was collected from each laboratory submitting a database and evaluated by the HUMN steering committee for compliance to acceptable standard methodology (10). Altogether, 20 laboratories from 10 countries, that were measuring MN frequency in human populations as a routine procedure, submitted databases that fulfilled the inclusion criteria, and all were included in the study.

Only individual MN frequencies based on the scoring of at least 1000 interphase cells were considered for statistical analysis. The large majority of laboratories adopted the cytokinesis-block assay (20), scoring MN frequency in binucleated lymphocytes cells. The dataset from Sweden was produced using a different protocol, and MN were scored in mononuclear lymphocytes (21), but this difference did not influence the statistical analysis of data.

The subjects were originally selected by the testing laboratories for cytogenetic *ad hoc* studies, or routine biological dosimetry because of their exposures to mutagens or carcinogens, or as unexposed referents. The original cytogenetic studies were performed between 1980 and 2002, and most of them have been published in the peer-reviewed scientific literature. The most commonly studied exposures were to ionizing radiation (1355 subjects), pesticides (318 subjects), polycyclic aromatic hydrocarbons (264 subjects) organic solvents (292 subjects) and cytostatic drugs (182 subjects). The cohorts investigated in this study were different from those that were previously studied to determine the relationship between CA in PBL and cancer risk (15–19).

All subjects included in the HUMN cohort had a valid personal identification code, were at least 15 years old, and free of cancer at the time of cytogenetic testing. The study protocol was approved by the ethics committee of the coordinating center at the National Cancer Research Institute of Genoa, Italy. The list of national cohorts with selected characteristics is reported in Table I. Overall, 6718 subjects, accounting for a total of 62 980 person-years, were studied.

Seventy subjects included in the cohort were tested more than once. Repeated measures showed a good degree of internal agreement ($r = 0.68$; $P < 0.01$), however only the result of the first test was considered for the statistical analysis. We used the 'non-concurrent' cohort study approach which is the most suitable epidemiologic design to evaluate long-term effects of early genetic damage, when the biomarker may be affected by the disease, i.e. reverse causality bias (22). The follow-up period began with the date of MN testing and ended with death, cancer diagnosis, emigration, 85th birthday or end of follow-up (1999–2004, depending on the country), whichever occurred first. The median duration of follow-up was 8.0 years. Information on cancer incidence was obtained by linking the cohorts with national or regional cancer registries. In Poland, an active system of follow-up was set up via contacts with local cancer registries, municipalities of residence, employers, pension funds and general practitioners. In Italy and Belgium, mortality data were available and information on causes of death was collected via postal follow-up from the Municipality of residence. It was possible to follow the incidence for a subset of 245 Italian subjects (7 cancer cases; 1459 person-years), and these data were included in the statistics of

**Table I.** Selected characteristics of the study population

| Country | Labs (N) | Subjects | Cancer cases | Person-years | Year(s) of test (min–max) | Median follow-up (years) | MN frequency [mean (SD)] | Age [years mean (SD)] | Males (%) | Exposed (%)[a] | Current smokers (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Belgium[b] | 1 | 146 | 0 | 599 | 1991–1996 | 4.0 | 3.9 (2.3) | 40.3 (9.3) | 100.0 | 50.0 | 45.9 |
| Bulgaria | 1 | 208 | 4 | 1264 | 1993–1996 | 6.3 | 41.1 (28.6) | 47.5 (8.5) | 63.0 | 67.3 | 57.7 |
| Croatia | 1 | 202 | 4 | 1841 | 1988–1999 | 9.5 | 59.0 (41.5) | 46.5 (9.9) | 71.3 | 100.0 | 53.0 |
| Italy[b,c] | 7 | 2654 | 56 | 25476 | 1984–1999 | 9.3 | 4.6 (4.2) | 54.9 (15.3) | 57.9 | 19.6 | 29.3 |
| Japan | 1 | 833 | 146 | 7935 | 1988–1990 | 10.6 | 52.2 (14.6) | 72.3 (9.6) | 40.8 | 100.0 | 26.6 |
| Poland | 1 | 205 | 11 | 1839 | 1989–1998 | 11.5 | 14.3 (10.0) | 50.8 (11.4) | 100.0 | 44.4 | 60.0 |
| Slovakia | 4 | 923 | 10 | 7735 | 1993–1998 | 8.0 | 12.1 (6.1) | 46.1 (9.6) | 39.1 | 72.6 | 38.3 |
| Sweden | 2 | 677 | 31 | 10440 | 1980–1986 | 15.8 | 4.1 (2.7) | 54.9 (11.4) | 87.7 | 59.0 | 40.3 |
| Taiwan | 1 | 194 | 2 | 1443 | 1983–1996 | 7.7 | 13.3 (11.1) | 42.2 (10.6) | 33.5 | 52.1 | 21.5 |
| Yugoslavia | 1 | 676 | 4 | 4408 | 1988–2002 | 5.6 | 11.6 (9.3) | 44.3 (9.0) | 53.7 | 69.2 | 41.1 |
| Total | 20 | 6718 | 268 | 62980 | 1980–2002 | 8.5 | 15.5 (21.0) | 53.7 (15.0) | 57.7 | 52.2 | 35.2 |

[a]Exposure to known or potential carcinogens at the time of blood sampling.
[b]Mortality follow-up.
[c]Incidence data were also available for 245 subjects out of 2654 (7 cancer cases; 1459 person-years)

cancer incidence. No major differences were found between incidence/mortality rates in the study cohorts and the national reference rates. Twenty-one subjects diagnosed with a non-melanoma skin cancer (ICD-IX 173) were excluded from the analyses.

In order to standardize for the marked inter-laboratory variability (see Table I) MN frequency was categorized by tertiles of the laboratory-specific distributions, i.e. ≤33%, 34–66%, >66%, and subjects were classified as having low, medium or high MN frequency, respectively. Information about occupational exposure to mutagens or carcinogens was collected as reported in the original studies, and re-classified according to the job-exposure matrix described by the ESCH group (17). Data on smoking status at the moment of cytogenetic testing were available in all cohorts, and subjects were classified as current, former or never smoker without consideration of the level of smoking. MN frequency was significantly increased in those subjects who were occupationally exposed to mutagens and carcinogens when compared with unexposed controls ($P < 0.005$), and to a lesser extent in heavy smokers when compared with those who never smoked ($P < 0.044$).

*Statistical methods*

Regression models were fitted to the data assuming that cancer incidence rates follow the negative binomial distribution (NB) (23). This distribution in presence of overdispersion, a phenomenon that frequently arises with count data, provides more efficient estimates of the standard errors of the models parameters. The NB variance function is

$$V(\mu) = \mu + \phi\mu^2$$

where $\mu$ is the NB mean and $\phi$ is the overdispersion parameter. NB reduces to the Poisson distribution when this parameter tends to zero.

The effect of MN frequency on cancer incidence was evaluated by comparing cancer incidence rates for the medium and high levels versus the low level, after adjusting for the confounding effects of age, gender, smoking status and occupational exposure to mutagens or carcinogens. A random effect term was then included in the models to adjust for the differences in cancer rates occurring among countries. The presence of effect modification was tested by computing the log-likelihood ratio test for two hierarchical models, the first with, and the second without interaction terms involving MN.

To verify if the presence of pre-clinical stages of cancer might have influenced MN frequency at test we repeated the analyses excluding the first 2 years of follow-up. Furthermore, we stratified the entire cohort according to the median follow-up time among the cancer cases, i.e. <6 and 6 years, and tested for the presence of effect modification due to time since test.

Given the homogeneous pattern of cancer incidence and cancer-free survival in the medium and high tertiles, the analyses by country and cancer site, which generally had a number of events too low to allow the analysis on three strata levels, were performed combining these two tertiles. These additional analyses were performed in groups with at least 10 observed cases.

The effect of MN frequency on the probability to be cancer free at the end of the follow-up was estimated by means of Cox's proportional hazard model (24), using time since test as the time variable and adjusting for age, gender, smoking status and occupational exposure. STATA software was used for all statistical analyses (25).

**Table II.** Relative risk of cancer incidence by MN frequency, gender, occupational exposure to carcinogens and smoking status

| Covariate | Cases | Subjects | Person-years | RR[a] | 95% CI | P-value |
|---|---|---|---|---|---|---|
| MN frequency | | | | | | |
| Low* | 51 | 1307 | 12415 | 1 | — | |
| Medium | 91 | 1426 | 13014 | 1.84[#] | 1.28–2.66 | 0.001 |
| High | 77 | 1430 | 12935 | 1.53[#] | 1.04–2.25 | 0.03 |
| Gender | | | | | | |
| Female* | 86 | 1814 | 15062 | 1 | — | |
| Male | 133 | 2349 | 23302 | 1.46 | 1.02–2.10 | 0.04 |
| Occupational exposure[b] | | | | | | |
| No* | 25 | 1169 | 10813 | 1 | — | |
| Yes | 194 | 2953 | 27274 | 1.41 | 0.85–2.35 | 0.19 |
| Smoking status[c] | | | | | | |
| Never* | 94 | 2272 | 20849 | 1 | — | |
| Former | 30 | 247 | 2246 | 1.23 | 0.76–2.01 | 0.40 |
| Current | 85 | 1575 | 14786 | 1.49 | 1.06–2.11 | 0.02 |

RR estimated by negative binomial regression analysis (4163 subjects).
All cancers: ICD IX 140–208.
[a]RR, relative risk adjusted also by age (5 years unit), country, time since MN test; CI, confidence interval.
[b]41 subjects with missing data.
[c]69 subjects with missing data.
*Reference group.
[#]Corresponding unadjusted RR (95% CI) for medium and high tertiles are 1.67 (1.20–2.34) and 1.47 (1.04–2.07), respectively.

## Results

An overall number of 219 incident cancers (including 7 cases from Italy) and 56 cancer deaths were registered at the end of the follow-up periods. The most frequent cancer sites were colon and rectum (37), stomach (35), lung (28) and breast (20), mostly contributed by Japan, Italy and Sweden where the person-years were the largest and the mean ages of the cohorts were the highest.

The association between overall cancer incidence and MN frequency is described in Table II. Significant increases were found for subjects in the medium group(Relative Risk (RR) = 1.84; 95% confidence interval (CI) 1.28–2.66) and high tertiles (RR = 1.53; 95% CI 1.04–2.25), when compared with the low tertile. Removing the first 2 years of follow-up (which most likely are affected by undiagnosed cancers)
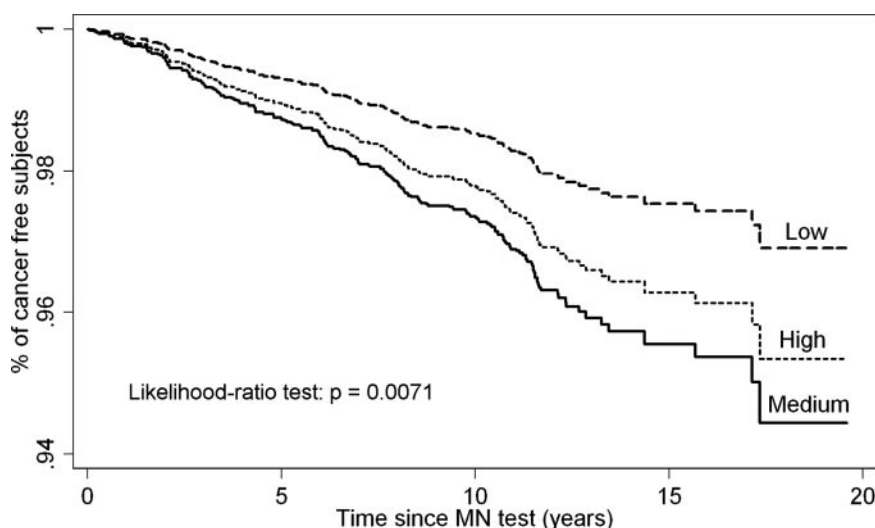
**Fig. 2.** Probability curves of cancer free survival by tertile of MN frequency (pooled data from the HUMN cohort). Cancer free survival refers to time from MN test to the first cancer diagnosis. The association between MN and survival was estimated using the multivariate Cox proportional hazard model. (all cancer incidence ICD IX 140–208).

**Table III.** Relative risk of cancer incidence by MN frequency and country

| Country | MN tertiles | Cases | Subjects | Person-years | RR[a] | 95% CI | P-value |
|---|---|---|---|---|---|---|---|
| Italy[b] | Low* | 12 | 812 | 7803 | 1 | — | |
| | Medium/high | 44 | 1842 | 17 673 | 1.37 | (0.72–2.62) | 0.34 |
| Japan | Low* | 42 | 274 | 2690 | 1 | — | |
| | Medium/high | 104 | 559 | 5245 | 1.53 | (1.06–2.21) | 0.02 |
| Sweden | Low* | 4 | 201 | 3287 | 1 | — | |
| | Medium/high | 27 | 476 | 7153 | 2.24 | (0.77–6.52) | 0.14 |
| Other countries | Low* | 4 | 754 | 5960 | 1 | — | |
| | Medium/high | 31 | 1654 | 12 570 | 2.85 | (0.99–8.23) | 0.053 |

RR estimated by negative binomial regression analysis (4163 subjects). All cancers: ICD IX 140–208.
[a]RRs adjusted by sex, age (5 years unit), time since MN test, exposure to carcinogens and smoking status at test.
[b]Mortality data.
*Reference group.

from the statistical analysis, did not change the risk estimates (13 cancer cases occurred within this time frame). Similarly, after splitting the follow-up period into two equal time periods, no differences were observed between RR's for medium and high levels of MN in the period immediately after the MN assay (<6 years) and in the period thereafter (≥6 years). The limited impact of confounding due to variables that affected MN frequency (i.e. age, gender, occupational exposure to genotoxins and smoking status) is shown by the small differences between the adjusted and unadjusted relative risks, i.e. 1.84 (1.28–2.66) versus 1.67 (1.20–2.34) for medium tertile and 1.53 (1.04–2.25) versus 1.47 (1.04–2.07) for the high tertile, respectively. However, after stratifying by gender, females showed a linear trend of RRs by MN frequency level, i.e. 1.00; 1.51; 1.96, for low, medium and high, respectively (test for linear trend, P < 0.03). Occupational exposure to mutagens or smoking status did not significantly modify the relationship between MN frequency and cancer risk.

The association between cancer risk and MN level is consistent over the follow-up time, as clearly shown by Figure 2, which indicated a significantly worse outcome in subjects in the medium and high MN tertiles (P = 0.001 and 0.025, respectively).

The MN frequency in the medium/high tertiles also presaged an increased risk of cancer within national cohorts, as reported in Table III. Only findings from Japan, which is the oldest national cohort and has the largest number of events, reached statistical significance, i.e. RR = 1.53 (1.06–2.21). However, a significant association (RR = 2.54; 1.19–5.40) between MN frequency and cancer risk was found when incidence data from all other countries (excluding Japan) were pooled.

The risks associated with specific cancer sites was tested after combining the medium/high tertiles and selecting those sites with a high number of cases (Table IV). All cancer sites evaluated—except for cancer of liver, biliary ducts and pancreas (RR = 0.63; 0.27–1.44)—showed a higher RR in the medium/high MN tertile. A significantly increased risk was found only for the group of urogenital cancers (ICD-IX 179–189), RR =2 .83; 1.19–6.74), with the highest RR being for bladder and kidney cancers (n = 16; RR = 8.23; 1.08–63.0). A significant association was also found by combining gastro-intestinal cancers (ICD-IX 151–154), RR = 1.74;

**Table IV.** Relative risk of cancer incidence and mortality by MN frequency and cancer site

| Tumour site (ICD IX) | MN tertile | Cases | RR[a] | 95% CI | P-value |
|---|---|---|---|---|---|
| Liver, biliary ducts and pancreas (155–157) | Low* | 14 | 1 | — | |
| | Medium/high | 15 | 0.63 | (0.27– 1.44) | 0.27 |
| Lung (162) | Low* | 7 | 1 | — | |
| | Medium/high | 21 | 1.54 | (0.63–3.77) | 0.35 |
| Stomach (151) | Low* | 8 | 1 | — | |
| | Medium/high | 27 | 1.63 | (0.70–3.82) | 0.26 |
| Colon-rectum (153–154) | Low* | 10 | 1 | — | |
| | Medium/high | 27 | 1.75 | (0.84–3.66) | 0.14 |
| Urogenital (179–189) | Low* | 6 | 1 | — | |
| | Medium/high | 36 | 2.83 | (1.19–6.74) | 0.002 |

RR estimated by negative binomial regression analysis.
[a]RRs adjusted by gender, age (5 years unit), time since MN test, exposure to carcinogens and smoking status at test.
*Reference group.

1.01–4.71. No statistics could be provided for breast cancer because none of the 14 incident cases fell in the reference tertile (7 in the medium and 7 in the high, respectively).

## Discussion

The results from the present study support the hypothesis that MN frequency in PBL is a predictive biomarker of cancer risk. These data strengthen and extend the evidence accumulated by experimental, mechanistic and association studies (1), that the extent of genetic damage measured in lymphocytes reflects the occurrence of early carcinogenic events in the target tissues. The strengths of this study include the relatively large size of the study group, the same direction of risk estimates in all countries, and the independence of results from the time elapsed between MN testing and cancer diagnosis. However there are some limitations, which are discussed below.

A specific association was found in a number of cancer sites, although statistical significance was reached only in the groups of urogenital and gastro-intestinal cancers. In particular, the higher risks for stomach (1.63) and intestinal cancers (1.75), (RR = 1.74; 1.01–4.71, when combined), are in agreement with the literature, which emphasizes the role of chromosome rearrangements in the early stages of these tumours (26–28). Further, a strong association between CA frequency in PBL and the risk of stomach cancer (n = 12; RR = 7.79; 1.01–60.0) has been described in a recent cohort study performed in the Czech republic (19), and in a new independent cohort of five Central and Eastern European countries (RR could not be evaluated because none of 15 cases of stomach cancer fell in the reference tertile, but the test for linear trend was highly significant; P < 0.01) (29). The lack of breast cancer cases in the low MN tertile relative to 14 breast cancer cases in the mid/high tertiles suggests that elevated MN may be related prospectively with breast cancer risk for which there is evidence of an association in case–control studies (30). Nevertheless, we acknowledge that the number of cancers per organ site is relatively small, and that the statistical estimates, which are suggestive of an association with MN, are likely to become more stable as further cancers accumulate with increasing age of the cohort.

A feature of this study is the non-linearity of the dose–response relationship between MN frequency and overall cancer incidence, showing that subjects in the medium and high tertile have a higher risk of cancer relative to the low tertile, but there was no significant difference between medium and high tertiles. A non-linear relationship between MN frequency and the risk of cancer seems the most likely explanation, assuming that there is a value of MN frequency beyond which no further increase in cancer risk occurs, e.g. cells with excessive genome damage may be eliminated by apoptosis. Alternative explanations include selective loss of genetic material, i.e. subjects with high frequency of MN might more easily lose chromosomes not often involved in carcinogenesis (such as sex chromosomes), or the role of nuclear budding, a mechanism to eliminate excess amplified DNA or chromosomes (4,31–34). Furthermore, genome damage-induced cell death has been found in many other diseases, such as neurodegenerative disease (35), which could also explain the plateau effect given that association with other degenerative disease was not explored. The presence of a linear trend of RRs in the subgroup of women, who were much less occupationally exposed relative to males (data not shown), suggests that the observed non-linearity may have, to some extent, been caused by incomplete adjustment for occupational exposure.

While a residual confounding due to occupational exposure to mutagens or smoking status is still possible (as shown by the small difference between adjusted and unadjusted relative risk estimates), statistical analysis showed a lack of effect modification, i.e. the risk associated with MN frequency is the same in exposed and unexposed, as well as in smokers and non-smokers. Similar results were described in a case–control study nested within the ESCH cohort study on CA (17). In that study the authors concluded that CA is a predictor of cancer risk which is independent from exposure to mutagen/carcinogenic agents (occupational exposure and smoking habit was re-assessed by an international group of occupational hygienists). The similarity of our findings seems to support the same conclusion in the present cohort study, but the limitation in the exposure assessment (original data as reported by authors were used on a yes–no basis) has limited the possibility of further evaluation of carcinogen/mutagen exposure effects.

Formation of nuclear anomalies such as MN, chromosomal rearrangements and anaphase bridges (leading to breakage-fusion-bridge cycles and generation of more MN) are events commonly seen in the early stages of carcinogenesis (4,26,36). Elevated levels of MN are indicative of defects in DNA repair and chromosome segregation which could

result in generation of daughter cells with altered gene dosage, or deregulation of gene expression that could lead to the evolution of the chromosome instability phenotype often seen in cancer (1,2,9,10,13,37). These considerations give mechanistic support to a possible causal association between MN frequency and the risk of cancer. The observed association between MN frequency and cancer risk in non-haematological malignancies in our study suggests that genome damage events in lymphocytes may be correlated with cancer initiating events in other tissues via a common genetic, dietary or environmental factor. This is also supported by a recent paper indicating that specific chromosomal rearrangements play a role not only in haematological malignancies but also in non-haematological malignancies, and that there may be no fundamental tissue-specific differences in the genetic mechanisms by which neoplasia is initiated (38).

A major challenge in the mechanistic interpretation of our findings is the fact that MN can be generated through different processes, i.e. chromosome breakage and chromosome loss (aneuploidy), occurring roughly in the same proportion (1). In contrast to chromosome breakage, whose role in early stages of carcinogenesis has extensively been studied, the significance of aneuploidy is still poorly understood, although it is well known that aneuploidy is a hallmark of the majority of human tumours, and is associated with high grade invasiveness and poor prognosis (39). Studies of the cytogenetic evolution in breast cancer have suggested that a highly aneuploid state could originate from a polyploidization event concurrent with a gradual loss of individual chromosome copies (40). Similar findings have also been reported for preneoplastic lesions of the colon (41), oesophagus (42) and cervix (43). Cancer risk estimates associated with MN frequencies in the present study are roughly similar to the estimates reported for CA. Our findings confirm the predictive role of chromosome anomalies but they do not provide definitive evidence concerning the role of aneuploidy. A resolution of this issue may result from future studies designed to distinguish between these mechanisms using centromere and telomere detection in MN, and including measurement of nucleoplasmic bridges which are a marker of misrepair of DNA lesions and telomere end-fusion (9,28,44).

The international multi-centre nature of this study may be considered among its strengths because this design allowed us to explore various environmental exposures in diverse genetic backgrounds (using the country as a proxy of genetic features). On the other hand, this approach has also introduced a number of limitations. These include the large inter-laboratory variability of MN frequency which is most likely due to technical differences in slide preparation and scoring, the heterogeneous quality of data on genotoxic exposures such as cigarette smoking and occupational carcinogens or the availability of a single measure of MN per individual which may have resulted in misclassification among MN frequency levels. A further potential source of bias is the heterogeneity in the cancer registration quality in the countries involved. All these potential sources of discrepancy—including the intrinsic limitation due to measurement of MN in a surrogate tissue—may have weakened the observed association between MN and cancer incidence. In addition, the young age of the cohort (mean = 53.7 years) limited the number of cancer cases, i.e. ~4% of the whole

study group, contributing to a reduced stability of risk estimates. The uneven contribution of national cohorts to the main database, and the much larger contribution of cancer cases from Japan, may have added some uncertainty to the random effects model and made the statistical estimates less stable. However, despite the uneven size of the cohorts, the increased RR's in all national cohorts, as shown in Table III, adds confidence to the reliability of the statistical model used for the overall analysis.

Additional research is needed not only to have a better insight into the association between the MN frequency and cancer, but also to evaluate the benefits of including biomarkers of cancer risk in the surveillance of populations at increased environmental or genetic risk. The first goal is easier to achieve, and plans already exist within the framework of the HUMN project for increasing the size of the study group, by both including new national cohorts and extending the length of the follow-up period for those cohorts currently included in the study. The second goal is more complex given the relatively low risk predicted by MN frequency and the limited evidence of specific association with cancer site or genotoxic agent. A better understanding of this association, especially taking into account the role of possible confounders and effect modifiers such as diet, oxidative stress and genetic polymorphisms, would be desirable before routine application of these biomarkers in population studies to estimate the risk of cancer.

Although the prospect of reducing chromosome damage and MN frequency by dietary, life-style and occupational changes may appear feasible (8,11,44), it will also be desirable and necessary to measure the actual impact of MN frequency reduction on cancer incidence prospectively.

In conclusion, this study provides preliminary evidence that MN frequency in PBL is predictive of cancer risk, suggesting that increased MN formation is associated with early events in carcinogenesis. However, it is important to emphasize that the findings of this study pertain to the risk of a group and not individuals. The wide-spread use of the MN assay in the monitoring of environmental and occupational exposure to genotoxins, its responsiveness to the effects of micronutrients and diet and its ability to identify high-risk groups of susceptible individuals, provides a further possibility for the use of the MN assay in the planning, implementation and validation of cancer surveillance and prevention policies.

## Acknowledgements

# References

1. Fenech,M., Holland,N., Chang,W.P., Zeiger,E. and Bonassi,S. (1999) The HUman MicroNucleus project—an international collaborative study on the use of the micronucleus technique for measuring DNA damage in humans. *Mutat. Res.*, **428**, 271–283.

2. Kimura,M., Umegaki,K., Higuchi,M., Thomas,P. and Fenech,M. (2004) Methylenetetrahydrofolate reductase C677T polymorphism, folic acid and riboflavin are important determinants of genome stability in cultured human lymphocytes. *J. Nutr.*, **134**, 48–56.

3. Bonassi,S., Ugolini,D., Kirsch-Volders,M., Strömberg,U., Vermeulen,R. and Tucker,J.D. (2005) Human population studies with cytogenetic biomarkers: review of the literature and future prospectives. *Environ. Mol. Mutagenesis*, **54**, 258–270.

4. Umegaki,K. and Fenech,M. (2000) Cytokinesis-block micronucleus assay in WIL2-NS cells: a sensitive system to detect chromosomal damage induced by reactive oxygen species and activated human neutrophils. *Mutagenesis*, **15**, 261–269.

5. Rajagopalan,H., Jallepalli,P.V., Rago,C., Velculescu,V.E., Kinzler,K.W., Vogelstein,B. and Lengauer,C. (2004) Inactivation of hCDC4 can cause chromosomal instability. *Nature*, **428**, 77–81.

6. MacGregor,J.T. (1990) Dietary factors affecting spontaneous chromosomal damage in man. *Prog. Clin. Biol. Res.*, **347**, 139–153.

7. Crott,J.W., Mashiyama,S.T., Ames,B.N. and Fenech,M. (2001) MTHFR C677T polymorphism does not alter folic acid deficiency-induced uracil incorporation into primary lymphocyte DNA *in vitro*. (Accelerated Paper). *Carcinogenesis*, **22**, 1019–1025.

8. Fenech,M., Baghurst,P., Luderer,W., Turner,J., Record,S., Ceppi,M. and Bonassi,S. (2005) Low intake of calcium, folate, nicotinic acid, vitamin E, retinol, β-carotene and high intake of pantothenic acid, biotin and riboflavin are significantly associated with increased genome instability—results from a dietary intake and micronucleus index survey in South Australia. *Carcinogenesis*, **26**, 991–999.

9. Fenech,M. (2002) Chromosomal biomarkers of genomic instability relevant to cancer. *Drug Discov. Today*, **22**, 1128–1137.

10. 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 lymphocyte cultures. *Mutat. Res.*, **534**, 65–75.

11. Ames,B.N. and Wakimoto,P. (2002) Are vitamin and mineral deficiencies a major cancer risk? *Nat. Rev. Cancer*, **2**, 694–704.

12. Van Schooten,F.J., Besarati Nia,A., De Flora,S. *et al.* (2002) Effects of oral administration of *N*-acetyl-L-cysteine: a multi-biomarker study in smokers. *Cancer Epidemiol. Biomarkers Prev.*, **11**, 167–175.

13. Chang,W.P., Hwang,B.F., Wang,D. and Wang,J.D. (1997) Cytogenetic effect of chronic low-dose, low-dose-rate gamma-radiation in residents of irradiated buildings. *Lancet*, **350**, 330–333.

14. Bettega,D., Calzolari,P., Doneda,L., Belloni,F., Tallone,L. and Redpath,J.L. (2003) Differential effectiveness of solar UVB subcomponents in causing cell death, oncogenic transformation and micronucleus induction in human hybrid cells. *Int. J. Radiat. Biol.*, **79**, 211–216.

15. Hagmar,L., Brøgger,A., Hansteen,I-L. *et al.* (1994) Cancer risk in humans predicted by increased levels of chromosomal aberrations in lymphocytes: Nordic Study Group on the Health Effects of Chromosome Damage. *Cancer Res.*, **54**, 2919–2922.

16. Hagmar,L., Bonassi,S., Strömberg,U., Brøgger,A., Knudsen,L., Norppa,H. and Reuterwall,C. (1998) Chromosomal aberrations in lymphocytes predict human cancer—a report from the European Study Group on Cytogenetic Biomarkers and Health (ESCH). *Cancer Res.*, **58**, 4117–4121.

17. Bonassi,S., Hagmar,L., Strömberg,U. *et al.* (2000) Chromosomal aberrations in lymphocytes predict human cancer independently from exposure to carcinogens. *Cancer Res.*, **60**, 1619–1625.

18. Hagmar,L., Stromberg,U., Bonassi,S., Hansteen,I-L., Knudsen,L.E., Lindholm,C. and Norppa,H. (2004) Impact of types of lymphocyte chromosomal aberrations on human cancer risk: results from Nordic and Italian cohorts. *Cancer Res.*, **64**, 2258–2263.

19. Rossner,P., Boffetta,P., Ceppi,M., Bonassi,S., Smerhovsky,Z., Landa,K., Juzova,D. and Sram,R.J. (2005) Chromosomal aberrations in lymphocytes of healthy subjects and risk of cancer. *Environ. Health Perspect.*, **113**, 517–520.

20. Fenech,M. and Morley,A.A. (1985) Measurement of micronuclei in human lymphocytes. *Mutat. Res.*, **147**, 29–36.

21. Högstedt,B. (1984) Micronuclei in lymphocytes with preserved cytoplasm. A method for assessment of cytogenetic damage in man. *Mutat. Res.*, **130**, 63–72.

22. Rothman,N., Stewart,W.F. and Schulte,P.A. (1995) Incorporating biomarkers into cancer epidemiology: a matrix of biomarker and study design categories. *Cancer Epidemiol. Biomarkers Prev.*, **4**, 301–311.

23. Lindsey,J.K. (1995) *Modelling Frequency and Count Data*. Oxford University Press, Oxford, UK.

24. Clayton,D. and Hills,M. (1993) *Statistical Models in Epidemiology*. Oxford University Press, Oxford, UK.

25. StataCorp. (2001) *Stata Statistical Software: Release 7.0*. STata Corporation, College Station, TX.

26. Norppa,H. (2004) Cytogenetic biomarkers. In: Buffler,P., Rice,J., Baan,R., Bird,M. and Boffetta,P., (eds) *Mechanisms of Carcinogenesis: Contributions of Molecular Epidemiology*. IARC Sci. Publ. Vol. 157, IARC, Lyon, pp. 179–205.

27. Stewenius,Y., Gorunova,L., Jonson,T., Larsson,N., Hoglund,M., Mandahl,N., Mertens,F., Mitelman,F. and Gisselsson,D. (2005) Structural and numerical chromosome changes in colon cancer develop through telomere-mediated anaphase bridges, not through mitotic multipolarity. *Proc. Natl Acad .Sci. USA*, **102**, 5541–5546.

28. Rudolph,K.L., Millard,M., Bosenberg,M.W. and DePinho,R.A. (2001) Telomere dysfunction and evolution of intestinal carcinoma in mice and humans. *Nat. Genet.*, **28**, 155–159.

29. Boffetta,P., van der Hel,O., Norppa,H. *et al.* (2006) Chromosomal aberrations and cancer risk: Results of a cohort study from Central Europe. *Am. J. Epidemiol.*, (in press).

30. Varga,D., Hoegel,J., Maier,C. *et al.* (2006) On the difference of micronucleus frequencies in peripheral blood lymphocytes between breast cancer patients and controls. *Mutagenesis*, Aug 22 [Epub ahead of print].

31. Wang,X., Thomas,P., Xue,J. and Fenech,M. (2004) Folate deficiency induces aneuploidy in human lymphocytes *in vitro*—evidence using cytokinesis-blocked cells and probes specific for chromosome 17 and 21. *Mutat. Res.*, **551**, 167–180.

32. Tanaka,T. and Shimizu,N. (2000) Induced detachment of acentric chromatin from mitotic chromosomes leads to their cytoplasmic localization at G(1) and the micronucleation by lamin reorganization at S phase. *J. Cell. Sci.*, **113**, 697–707.

33. Decordier,I., Cundari,E. and Kirsch-Volders,M. (2005) Influence of caspase activity on micronuclei detection: a possible role for caspase-3 in micronucleation. *Mutagenesis*, **20**, 173–179.

34. Decordier,I., Dillen,L., Cundari,E. and Kirsch-Volders,M. (2002) Elimination of micronucleated cells by apoptosis after treatment with inhibitors of microtubules. *Mutagenesis*, **17**, 337–344.

35. Migliore,L., Petrozzi,L., Lucetti,C. *et al.* (2002) Oxidative damage and cytogenetic analysis in leukocytes of Parkinson's disease patients. *Neurology*, **58**, 1809–1815.

36. Gisselsson,D. and Hoglund,M. (2005) Connecting mitotic instability and chromosome aberrations in cancer—can telomeres bridge the gap? *Semin. Cancer Biol.*, **15**, 13–23.

37. Kirsch-Volders,M., Sofuni,T., Aardema,M. *et al.* (2003) Report from the *in vitro* micronucleus assay working group. *Mutat. Res.*, **540**, 153–163.

38. Mitelman,F.., Johansson,B. and Mertens,F. (2004) Fusion genes and rearranged genes as a linear function of chromosome aberrations in cancer. *Nat. Genet.*, **36**, 331–334.

39. Kops,G.J., Weaver,B.A. and Cleveland,D.W. (2005) On the road to cancer: aneuploidy and the mitotic checkpoint. *Nat. Rev. Cancer*, **10**, 773–785.

40. Lingle,W.L., Barrett,S.L., Negron,V.C., D'Assoro,A.B., Boeneman,K., Liu,W., Whitehead,C.M., Reynolds,C. and Salisbury,J.L. (2002) Centrosome amplification drives chromosomal instability in breast tumor development. *Proc. Natl Acad. Sci. USA*, **99**, 1978–1983.

41. Cardoso,J., Molenaar,L., de Menezes,R.X., van Leerdam,M., Rosenberg,C., Moslein,G., Sampson,J., Morreau,H., Boer,J.M. and Fodde,R. (2006) Chromosomal instability in MYH- and APC-mutant adenomatous polyps. *Cancer Res.*, **66**, 2514–2519.

42. Doak,S.H., Jenkins,G.J.S., Parry,E.M., D'Souza,F.R., Griffiths,A.P., Toffazal,N., Shah,V., Baxter,J.N. and Parry,J.M. (2003) Chromosome 4 hyperploidy represents an early genetic aberration in premalignant Barrett's oesophagus. *Gut*, **52**, 623–628.

43. Olaharski,A., Sotelo,R., Solorza-Luna,G., Gonsebatt,M.E., Guzman,P., Mohar,A. and Eastmond,D.A. (2006) Tetraploidy and chromosomal instability are early events during cervical carcinogenesis. *Carcinogenesis*, **27**, 337–343.

44. Fenech,M. (2005) The Genome health clinic and genome health nutrigenomics concepts: diagnosis and nutritional treatment of genome and epigenome damage on an individual basis. *Mutagenesis*, **20**, 255–269.