

Ministerial Advisory Group on Veterans' Health  
PO Box 5146, Wellington

23 December 2010

Hon Judith Collins  
Minister of Veterans Affairs  
Parliament Buildings  
Wellington



Dear Minister

The Ministerial Advisory Group on Veterans Health has reviewed the research undertaken by Massey University on behalf of the New Zealand Nuclear Test Veterans Association (NZNTVA).

We submitted the cytogenetic studies completed at Massey University by Professor Al Rowland and colleagues, and the psychological studies completed by Associate Professor John Podd and colleagues to national and international expert external review.

The Advisory Group has evaluated the research, considered the reviews and has sought responses to the reviews from the Massey University researchers and from the NZNTVA.

The Massey University studies are observational, case control studies. In such studies the validity of the results is very dependent on how well the cases are matched with the controls, the size of the sample and how well the sample represents the group as a whole. If the cases, in these studies the nuclear test veterans, are not well matched with the controls then any difference in findings may be attributed to differences other than the nuclear test exposure.

I have enclosed the detailed reports we have received. I have also enclosed an earlier report from Professor Stephen Robertson, who reviewed the transgenerational implications of the Massey University chromosomal studies, an issue of particular importance to the veterans.

## CYTOGENETIC STUDIES

The cytogenetic studies were completed on 50 nuclear test veterans and 50 controls who were ex-servicemen, mainly from the army and some police. The rate of smoking was higher in the nuclear test veterans and smoking does cause chromosomal mutations. The researchers have attempted to control for this.

A number of different methodologies were used to assess chromosomal damage. The results from the multicolour fluorescent in situ hybridisation (mFISH) assay are the only results the Advisory Group considered reliable in determining past radiation exposure.

The mFISH study did demonstrate an increased number of stable translocations in the nuclear test veterans. From these results it is reasonable to conclude that the nuclear test veterans were exposed to ionising radiation but it is not possible to determine the exposure dose.

The clinical effects of this radiation exposure and the chromosomal changes are unknown. The only way of determining whether the health of the nuclear test veterans has been affected by their exposure would be through a morbidity and mortality study. Funding for such a study is available if the veterans wish to participate.

In conclusion:

i) the Massey University mFISH study results do provide evidence that the nuclear test veterans were exposed to ionising radiation. It is not possible to determine the extent of the exposure from these studies.

ii) the clinical consequences of this, if any, are not known.

iii) there is no reason to believe that these changes would have a transgenerational effect and have adverse health consequences for the children and grandchildren of the veterans. The Advisory Group found no evidence in the literature that children of nuclear veterans were at increased risk of inherited disorders.

iv) the Advisory Group has no evidence to support the addition of further conditions to the current list of conditions presumptively accepted as being service related for New Zealand nuclear test veterans.

## PSYCHOLOGICAL IMPACT STUDIES.

The psychological impact studies used the same nuclear test veterans and controls as the cytogenetic studies. There were a number of exclusion criteria necessary for the cytogenetic studies. Unfortunately these exclusion criteria and the method of case and control selection have seriously limited the conclusions which can be drawn from the psychological impact studies.

Because of the method of selection of the veterans and the response rate the Advisory Group was concerned that

- i) the sample of veterans studied was not representative of all nuclear test veterans,
- ii) the control group had better mental health than a group of similar age and sex in the general population,
- iii) the veterans and the controls differed in a number of ways other than in their exposure to nuclear radiation and these differences may have accounted for the psychological differences between the two groups.

Also, the differences between the two groups were in mean scores on depression and quality of life assessment instruments. These differences were statistically significant but this does not mean that they were of clinical importance or represented an increased prevalence of psychiatric illness.

From our meetings with the NZNTVA and our review of the Massey University psychological impact studies the Advisory Group became very aware of the stress that has been caused to many of the nuclear test veterans and their families by the international failure to acknowledge that they were put in harm's way during nuclear testing and by their continued concern about the health implications of their exposure to ionising radiation. Interpreting the research is difficult because of the technical nature of the studies, the length of time since exposure and the flawed design of many of the studies. Poorly designed and reported studies fuel the concern. We were very conscious of the concerns veterans had for the health of their children and grandchildren when there is really no evidence from the genetic studies that there is any increased risk of inherited problems. The Advisory Group considers that there are steps New Zealand could take to address these issues.

## RECOMMENDATIONS.

i) the Advisory Group recommends no change to the current list of conditions presumptively accepted as being service related for New Zealand nuclear test veterans.

ii) The Advisory Group recommends that the Government acknowledges that the nuclear test veterans were put at risk through exposure to nuclear radiation, and that we have been slow to address the concerns of the veterans. There have been some suggestions as to how this might be done and I would like to discuss these with you.

iii) the scientific literature is technically difficult and often methodologically flawed. Poor studies have led to a fear of health consequences for veterans and their families when there is no real medical basis for such concerns. We recommend that we commission a critical review of the literature to be written for a lay readership by an author experienced in communication under the guidance of the Advisory Group. The purposes of such a document would be to provide a clear and unbiased assessment

of the consequences of nuclear test exposure and to help allay the unnecessary health concerns of many nuclear test veterans and their families.

I would be pleased to discuss these recommendations with you.

Yours sincerely,

A handwritten signature in blue ink, appearing to read "John Campbell", with a horizontal line underneath the name.

Professor John Campbell  
Chairman  
Ministerial Advisory Group on Veterans' Health



## Review of New Zealand Nuclear Test Veterans' Study – A Pilot Project (Psychological Impact)

This project seeks to examine whether there are any effects on psychological indices and well-being of a group of NZ veterans exposed to nuclear radiation compared to a control population. The study assesses personnel who took part in Operation Grapple in the Pacific during 1957-58.

### Methodology

In order to establish whether differences exist in the exposed population the ideal methodology would be to choose a random sample of exposed individuals and compare them to a random sample of naval personnel serving at the same time who were not involved in Operation Grapple. Deviation from this approach opens the possibility of bias influencing the results and limiting the interpretations that can be drawn from the data.

Looking first at the exposed group, the weakness of the current design is that the exposed group were self-selected (not randomly chosen) from recruitment by the NZ Nuclear Test Veterans' Association, a local pressure group involved in seeking recognition for health claims from exposure. This does open the possibility of error through veterans with health problems being more likely to volunteer for the study and being more focused on negative health effects because of this belief [1,2].

The control group was recruited through the RSA and through *personal contacts of the exposed participants*. This also opens the opportunity for error through exposed participants consciously or unconsciously choosing controls that may highlight differences in health status, given that the purpose of the study was clear. Furthermore, rather than choosing similar aged unexposed Naval personnel, the controls were drawn from the NZ Police, Army or from men who had "some form of compulsory military training". This opens the possibility that the control group may differ from the exposed group in more ways than their history of radiation exposure. Indeed, while the control is well matched for age, it does seem to be a more educated and higher income group.

Clearly the investigators may have been limited by budgetary factors and access to Naval records but the methodology used has created a major limitation, as it has opened the possibility that the exposed and control groups do not fairly represent the populations of interest.

### Analysis

The analysis of the data seems very competently carried out. Overall, the results show that the exposed sample report lower levels of well-being, higher levels of depression and more memory problems. The exposed group report a greater number of health problems but the significance of these is hard to determine due to the small numbers in the study.

I compared the exposed men's scores on the SF-36 with the norms for this scale for NZ males aged 65 to 74 years [3]. The mean of the SF-36 NZ norm scores falls roughly between the exposed and control groups and the exposed group again scores lower on each SF-36 dimension. This does also suggest that the control group may be in better health than a similar aged group in the general population.

### Interpretation

The weakness of the design limit the interpretation of the differences between the exposed and control group, as exposure to radiation is unfortunately probably not the only way in which these groups differ. It also still leaves open the possibility that the differences were caused by radiation in the late 1950's.

There are other possibilities that may explain the differences. The authors of the report suggest that the exposed men are suffering from chronic stress due to their beliefs about the danger of the radiation they were exposed to during their time in the Navy. This interpretation is difficult to evaluate since stress levels were not measured in the survey. It would be also necessary to establish how often men thought or worried about their radiation exposure in order to sustain this argument.

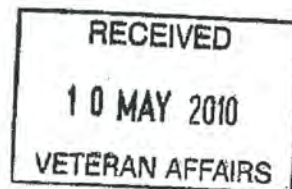
Another more probable explanation, especially given the way the sample was collected, is that the belief that exposure to radiation has caused harm to their body or health has in itself caused an increase in depression and physical symptoms as well as decreased well-being [4,5]. The authors of the report do themselves mention this possibility on page 43. Recent examples of this process are Gulf War Illness and the effects of the Amsterdam Air Disaster (where a belief existed there was uranium in the plane cargo) on civilian and emergency workers' health [2,6,7].

### Conclusions

The weakness of this study stem the methodology in selecting the samples for the exposed and control groups. The study itself and the analysis are very competently carried out. The data obtained show a clear decrement on psychological and well-being measures in the exposed group compared to the control group. The interpretation of these differences as reflecting differential radiation exposure between the groups is difficult to sustain given the biases in samples selection. The selection of the samples suggests another more possible explanation of the physical and mental health problems reported by exposed participants being caused by a belief that radiation has caused harm to their health and well-being. Without a study that evaluates the health of random samples of exposed and non-exposed Naval personnel, the question of negative health effects from Operation Grapple remains open.

---

Keith J Petrie PhD  
Professor of Health Psychology  
Faculty of Medical and Health Sciences  
University of Auckland



## References

1. Petrie, K.J., Broadbent, E., Kley, N., Moss-Morris, R., Horne, R., & Rief, W. (2005). Worries about modernity predict symptom complaints following environmental spraying. *Psychosomatic Medicine*, 67, 778-782.
2. Stuart, J.A., Ursano, R.J., Fullerton, C.S., Norwood, A.E. & Murray, K. (2003). Belief in exposure to terrorist agents: Reported exposure to nerve or mustard gas by Gulf War veterans. *Journal of Nervous and Mental Disease*, 191, 431-436.
3. Ministry of Health. (1999). *Taking the pulse: The 1996/97 New Zealand Health Survey*. Ministry of Health: Wellington, N.Z.
4. Jones, E., Palmer, I., & Wessely, S. (2007). Enduring beliefs about effects of gassing in the war: qualitative study. *British Medical Journal*, 335, 313-5.
5. Page, L., Petrie, K.J., & Wessely, S. (2006). Psychosocial responses to environmental incidents: A review and proposed typology. *Journal of Psychosomatic Research*, 60, 413-422.
6. Huizink, A.C., Slottje, P., Witteveen, A.B. (2006). Long-term health complaints following the the Amsterdam Air disaster in police and fire-fighters. *Occupational and Environmental Medicine*, 63, 657-62.
7. Iverson, A., Chalder, T., & Wessely, S. (2007). Gulf War Illness: lessons from medically unexplained symptoms. *Clinical Psychology Review*, 27, 842-854.

## Report for Ministerial Advisory Group on Veteran's Health

Ian Morison 14 July 2010

I have received two reports from the New Zealand Nuclear Test Veteran's Study. The first is entitled "New Zealand Nuclear Test Veteran's Study – A Pilot Project (Sister Chromatid Exchange)", authored by Rowland RE, Podd JV, Wahab M from the Institute of Molecular BioSciences, and the School of Psychology, Massey University.

The second report, "New Zealand Nuclear Test Veteran's Study – a Cytogenetic Analysis" was authored by Rowland RE, Podd JV, Wahab MA, Nickless EM, Parmentier C and M'Kacher R from the Institute of Molecular BioSciences and The School of Psychology, Massey University and Institut Gustave-Roussy, France.

### Report 1: "New Zealand Nuclear Test Veteran's Study – A Pilot Project (Sister Chromatid Exchange)"

I found this report disappointing for several reasons.

Reading the introduction leads one to believe that Sister Chromatid Exchange (SCE) testing is an accepted methodology in the area of biodosimetry following radiation exposure. However, the SCE test is not mentioned in recent reviews on cytogenetic methods for biological dosimetry (for example (Leonard *et al.* 2005; Kleinerman *et al.* 2006; Simon *et al.* 2010). Amongst the references cited by the authors on page 5, only two (Prabhavathi *et al.* 1995; Lazutka *et al.* 1999) describe the results for SCE *in vivo*; the other references refer to *in vitro* exposure.

Although they report an increased rate of SCE in Chernobyl clean up workers, Lazutka *et al.* comment that "*increased frequencies of SCE in Chernobyl clean up workers were also unexpected, since it is well known that SCE are not a very sensitive indicator of exposure to ionising radiation.*" (Lazutka *et al.* 1999).

The introduction appears to be biased towards presenting material that supports the utility of SCE and that supports the hypothesis that the veterans have a significant risk of ill health subsequent to exposure to radiation. For example, Rabbitt Roff's report of an increased frequency of myeloma is cited, but they fail to give weight to Muirhead *et al.*'s conclusion "*there was no evidence of an increased raised risk of multiple myeloma among test participants in recent years, and the suggestion and the first analysis of this cohort of a raised myeloma risk relative to controls is likely to have been a chance finding.*" (Muirhead *et al.* 2003).

They devote half a page (page 5) justifying the use of Sister Chromatid Exchange, but present no data or discussion that it might have limited applicability. I found the sentence "*any damage to DNA is universally accepted as being detrimental to a person's well being.*" excessive.

#### Methods Section

The methods section implies that all participants were recruited at the same time. However, Information Sheet (c) states: "*we have already obtained a group of veteran's who were exposed to a nuclear bomb blast. What we now need is a comparison group who are very similar to these men, but who were not so exposed.*" Given that the Information Sheet would normally be provided immediately prior to consent, and that the consent form related to providing a sample of blood, I am concerned that the blood samples of the controls were collected at a different time than the cases.



Given that lymphocyte culture and SCE tests can vary with the batches of reagents, I would be concerned about potential bias. (See additional comments on Report 2).

The methods section gives no information about the statistical analysis (see results).

### Results

A plethora of statistical tests are provided but there are insufficient data by which to judge them. First, the results are presented in terms of mean and confidence intervals, and then as a t-test. This analysis assumes that the independent unit of analysis is the cell rather than the participant. That is, the confidence intervals and t-test results are based on 2057 and 1635 cells respectively. However, the design of the study is a nested design in that 50 exposed subjects are compared to 50 controls. Figure 4 clearly shows that some individuals have low Sister Chromatid Exchange values and some have high. Therefore it is almost certain that the results for the 50 cells from a single individual are likely to be more closely correlated than those between individuals. In view of this nested design, the t-test result and the confidence intervals are incorrect.

The authors then switch to a non-parametric test based on the median result for each subject. Although this approach is more conservative, the significance test is not corrected for cigarette smoking, a factor that was much higher in the exposed subjects.

Assuming that the assay has some biological relevance, it would seem that the key analysis is the rate of Sister Chromatid Exchange using an appropriate nested design, adjusted for smoking. The sentence on page 16 that describes this analysis is inadequate.

I am concerned about the presentation of Figures 3 and 4. From Figure 4 it is not apparent that any of the cases are outliers with respect to SCE, yet in Figure 3, so called outliers have been excluded apparently on the basis of their SCE result. It does appear, however, that these excluded cases are outliers with respect to the 'Proportion'. There seems to be no good reason to exclude these cases.

### Conclusion

There is very little published data on the use of Sister Chromatid Exchange analysis in subjects exposed to radiation. From my reading, there appears to be no support for the use of this test in this setting. Therefore, I believe there is no basis for using this test to determine whether New Zealand nuclear test veterans had significant exposure to radiation. I am disappointed by the analysis, and without access to primary data and the services of an expert statistician, it is impossible to determine whether the reported differences were indeed significant. My hunch is that with an appropriate nested design and appropriate adjustment for smoking, there would be no significant differences between the two groups.

## Report 2: "New Zealand Nuclear Test Veteran's Study – a Cytogenetic Analysis"

### Introduction

The introduction of this report is similar to those for the report on Sister Chromatid Exchange. For example the report of an increased frequency of myeloma in British veterans is cited but no emphasis is given to papers that showed no increase in cancer.

The discussion of the G2 assay and the Micronucleus assay include no discussion of the lack of utility of these assays for retrospective biodosimetry.

The G2 assay measures the ability of cells to repair DNA damage induced by acute radiation exposure. It is not a recognised measure of past radiation exposure. Although it is possible that exposure to radiation in the distant past caused mutations of DNA repair genes, it is unlikely that

exposed veterans would have specific mutations of these genes. Of interest, recently Hamasaki *et al.* (2009) showed no evidence of cytogenetic instability in clonal T cells expanded in vitro for 25 population doublings. I am not aware of any literature that supports the role of DNA repair defects in biodosimetry.

Similarly the Micronucleus assay is a recognised marker for measuring acute radiation damage. To quote Leonard *et al.* (2005) "the poor resolving power of this assay, the inter-laboratory variability of dose-effect relationships for an acute low-LET radiation and the decrease with time of the anomalies explain why the micronucleus assay has been rarely used in biological dosimetry." In Table 1 they state that the usefulness of the micronuclei assay for retrospective biological dosimetry is "None".

On page 6 of the introduction there is the discussion of Hande *et al.*'s (2003) results, but most of this section, relates to exposure to inhaled plutonium. No evidence is presented that inhalation of plutonium or other radioactive material is relevant to these veterans.

Presently the most suitable biodosimetry methods for epidemiological studies are chromosome aberration frequencies from fluorescent in situ hybridisation (FISH) of peripheral blood lymphocytes and electron paramagnetic resonance (EPR) measurements made on tooth enamel (Leonard *et al.* 2005) (Kleinerman *et al.* 2006) (Simon *et al.* 2010). I am disappointed that this study did not focus more carefully on assays that have been validated for biodosimetry.

#### Methods Section:

I have the same concern as with the previous report, that the control group was sampled after the exposed group. Based on the participant identification numbers, it appears that the majority of exposed subjects were recruited in the first half of the study and the majority of controls in the second half. Bender *et al.* were concerned about the consistency of culture conditions stating: "because of the importance of time of sampling and proliferation rate in culture, any factors which influence the rate of cell progression in culture may influence the measurement of aberration frequencies. Blood cells from different individuals may respond differently to the stimulating effect of PHA, and different culture media and/or sera also give different cell progression rates." (Bender *et al.* 1988)

However, I have compared the rate of translocations against the participant identification number and it is pleasing that the period of recruitment does not appear to affect the number of translocations per cell (my Figure 1).

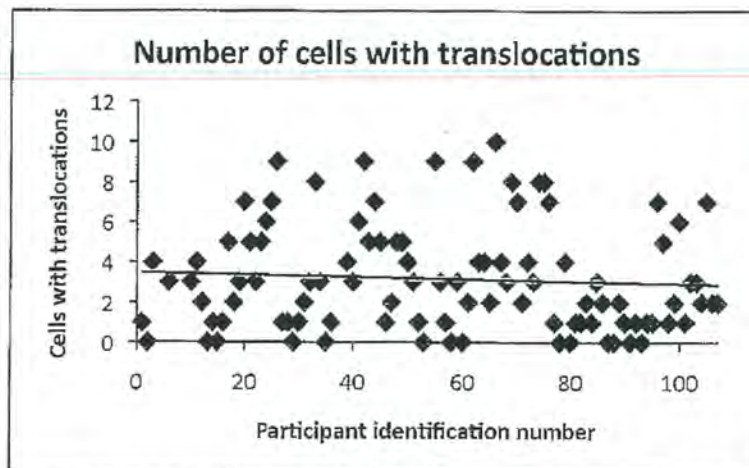


Figure 1. The time of recruitment did not appear to affect the number of translocations. These results include exposed and control participants.

## Results

### ***G2 assay and Micronucleus assay***

As expected the G2 assay and the Micronucleus assay showed no differences between the exposed and the control groups. Since neither assay is relevant to the research question, they will not be discussed further.

### ***Dose reconstruction***

Although the dose reconstruction section is approached with some caution, I perceive from my reading of the biodosimetry literature that there is no support for such an approach. I have not considered it any further.

### ***Multicolour fluorescent in situ hybridisation (mFISH)***

The most difficult part of this report relates to the results from mFISH. The mFISH assay is the only cytogenetic assay recommended in the literature for use in biodosimetry for past radiation exposure.

I wish to consider three aspects of the mFISH results:

1. Unstable chromosome aberrations;
2. Very complex cells;
3. Stable translocations.

#### ***1. Unstable chromosome aberrations***

As noted in the discussion (page 40), unstable chromosome aberrations such as dicentric chromosomes are not useful for distant biodosimetry, but despite this knowledge, the authors draw some strong conclusions stating: *"This high frequency of dicentric chromosomes in the veterans, including those that were observed but not scored in very complex cells, is very evocative of irradiation and suggests to us that the veterans may have been contaminated and may have retained high-LET long half-life radionuclides in their bodies."*

The literature is quite clear that unstable chromosomal aberrations disappear from the circulation during a few years following exposure. McLean and Michie (1995) explain this concept clearly: *"Cells that sustain unstable chromosomal lesions during irradiation die in their next mitosis. The rate at which lymphocytes with such damage disappear from the circulation, therefore, gives an estimate of the sum of their death rate and their division rate."*

Leonard et al. comment: *"The technique [scoring the frequency of dicentric chromosomes] is expensive, requires well-trained personnel and is unsuitable when a period of years has elapsed after exposure because of the limited lifespan of unstable aberrations."* In Table 1 they state that the usefulness of quantification of the unstable dicentric chromosomes for retrospective biological dosimetry is *"None"* (Leonard et al. 2005).

Bender et al. (1988) show the expected decline in the prevalence of unstable chromosomal aberrations, showing previous work by Evans et al. from patients irradiated for ankylosing spondylitis. After about 20 years the level of unstable aberrations reduces to apparently background levels.

Although not discussed in the text of the results section, the number of unstable chromosomal aberrations is provided in Tables 10 and 11, and in the Appendices. In the current study the exposed participants show a much higher rate of dicentric chromosomes than controls. However, the rate in the exposed participants (1.3 per 1000 cells) is comparable to the mean and median of 16 studies of control subjects summarised by Bender et al. (1988). In their Table 6.1, Bender et al. show the

frequency of dicentric chromosomes in control subjects from 16 studies. The median was 1.4 per 1000 cells (range 0-2.8) and the mean 1.3 per 1000 cells. I would therefore conclude that the low rate of dicentric aberrations in the control participants was the unexpected result.

Given the difficulties in quantifying the number of dicentric chromosomes because of their occurrence in complex cells, and given that the rate in exposed participants is the same as the mean of these 16 studies, I am not convinced by Rowland *et al.*'s view, reiterated in summary, that the difference is meaningful.

## 2. Very complex cells (rogue cells).

Although not mentioned in the results section and not quantified in the appendices, in the discussion the authors refer to their observations that the number of extraordinarily complex cells ('rogue cells') "observed in the Control group amongst the thousands of cells observed, amounted to less than 10. In contrast, the number of rogue cells observed in the Experimental group (but not scored) amounted to a few hundred." (p. 41). I am disturbed by this comment, since whilst they have "no precise record of their frequency" (p. 41), they obviously made these observations in a study in which "The codes were broken and Experimentals/Controls identified only after all genetic analyses were completed" (p. 10). It seems to me that their statements are incompatible and I am deeply concerned that their unblinded observations (e.g., the high number of unrecorded dicentrics; p. 40) have influenced their comments in the discussion section.

There is some discussion of 'rogue cells' in the literature and these concerned me. For example, Bender *et al.* state "Another problem area is the handling of high aberration-frequency cells, sometimes called "rogue" cells, that usually contain several chromosome - type exchanges and a wide range (2-> 10) of interstitial deletions ("double minutes"). The frequency of these cells varies from individual to individual but, in those studies where they have been observed, is generally in the range of 1 in 2000-5000 cells (Awa and Neel, 1986). In the large study of Brookhaven national Laboratory employees (Bender *et al.*, 1986), no "rogue" cells have yet been observed in a sample approaching 100 000 cells from 500 subjects. The presence of such high-aberration-frequency cells does not appear to be related to a known exposure to a clastogen, and at present the mechanism whereby these multiple aberrations arise is not known. This makes the significance of such cells rather equivocal. However, their contribution to background frequencies of aberrations could be very significant." (Bender *et al.* 1988)

In addition, Lazutka *et al.* comment "Forty rogue cells were found among 509 Chernobyl clean-up workers and 50 controls. However, the frequency of rogue cells was very strongly scorer-dependent: 31 individuals with rogue cells among 210 individuals analyzed were found by one scorer, while only nine individuals with rogue cells among 349 individuals analyzed were found by other three scorers." (Lazutka *et al.* 1999).

Although only mentioned in the discussion, I do not believe there is a sound basis for including the number of 'rogue cells' in any overall interpretation of the data presented in this report.

## 3. Stable translocations

Stable translocations have the potential to persist for decades after exposure to radiation and are generally accepted as being among the most useful of the biodosimetry markers. It must also be noted that the utility of this diagnostic test has not been well evaluated. "The kinetics of translocation loss in humans has not been well characterized, because there have been relatively few opportunities to perform such studies." (Kleinerman *et al.* 2006).

The New Zealand War Veteran's Study reports an increased rate of stable translocation among the exposed participants (p.26). Unlike the Sister Chromatid Exchange report, the statistical analysis

seems intuitively correct. From my reading of the literature I see no obvious reason to doubt the fundamental finding. The differences are large and obvious (my Figure 2).

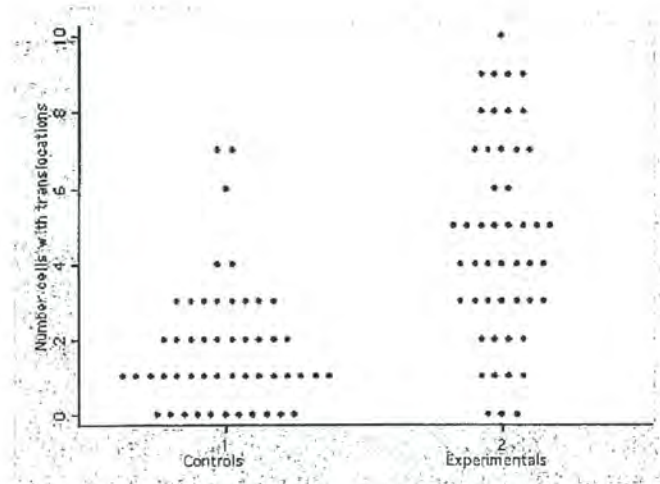


Figure 2. Number of cells with translocations (uncorrected for number of cells counted)

There are some outstanding questions that are not well addressed.

First, what confidence do we have that the selected cells (the non-complex, stable translocation subset) are a distinct subset of cells that reflect distant past exposure? For example, as shown in Appendix IV, the majority of scored abnormal cells have abnormalities other than stable translocations. In the control group, these abnormal, non-stable cells were five times more frequent than the 'stable' cells, and in the experimental group, these cells are 2.5 times more frequent. In addition it appears that none of the complex cells, and certainly none of the rogue cells were included in these counts. We do know that the number of scored dicentric and acentric cells was "a gross underestimate" because of the large number of complex cells (p. 29). Therefore, we cannot know what proportion of abnormal cells have 'stable' translocations. If the proportion was small, as it appears to be, one wonders whether they can be specifically selected to reflect past exposure. Alternatively, are they merely an undefined subset from a mixture of aberrant cells induced by lymphocyte culture?

Second, how well has smoking been considered as a confounding variable? The analysis on page 30, dichotomises by ever vs. never smokers, yet we know that the experimental smokers smoked on average about 1.8-fold more units than the control smokers. Given that smokers had lower translocation scores than never-smokers, it does seem unlikely that a more sophisticated analysis based on quantity of units smoked would alter the conclusion. It is of some concern, that the lack of effect of smoking contrasts with some of the literature, as Kleinerman *et al.* discuss: modifiers of translocation frequency include "subject's age and smoking status, because both have been significantly associated with increases in translocation frequencies in almost every study with sufficient statistical power. Translocations can be viewed as a dosimeter that is the integral of all of life's exposures (e.g. chemical, viral, environmental), including but not limited to radiation." (Kleinerman *et al.* 2006). In contrast, as stated in Report 2, Whitehouse *et al.* found no evidence that smoking increases the translocation rate in controls (Whitehouse *et al.* 2005).

### Conclusions

I recommend that the results from the sister chromatid exchange assay, the G2 assay, the Micronucleus assay, the dicentric cells and the rogue cells, and the dose reconstruction studies be excluded from consideration. None of these tests are recommended for the evaluation of distant

past radiation exposure. In addition, the results from these tests were either negative or unconvincing.

The difference between exposed and control subjects with respect to stable translocations do appear to be robust and is supported by the literature. I have some concerns as noted above, especially the question of what proportion of abnormal cells were scored, but I see no fundamental reason to discount the finding.

It is important to put the findings of Rowland *et al.*'s report in the context of the international review literature, which is quoted:

From Leonard *et al.* (2005):

*"In conclusion, biological dosimetry has serious limitations exactly for situations where the need for information is most urgent. It renders its most useful results when an individual has been exposed to a rather homogeneous high-level radiation over a short time interval, i.e. accidents at high-intensity radiation devices. On the other hand, it yielded less satisfactory information even when the most recent techniques were used for situations, where a low level, low dose rate exposure has occurred at some time in the past, for example for persons living in areas contaminated from the Chernobyl accident. Such negative experiences should be kept in mind in order to avoid futile and expensive investigations in the case of populations exposed from radioactivity and, notably, also from potentially clastogenic chemical agents."*

*"Regarding retrospective dosimetry years after an exposure or when a population continues to be exposed to a doubtful radiation or chemical risk, it will often be difficult to decide whether it is worthwhile to initiate a large scale bio-monitoring of populations for radiation or chemical exposure. Thereby, one must carefully consider costs and utility. Tens of thousands of cells may have to be studied in several hundred people in order find out more than the simple fact that a population has been exposed to some radiation or genotoxic chemical. Moreover, it is frequently impossible to take account of confounding factors in such studies. Consequently, such bio monitoring often has caused more confusion and anxiety in the population than has yielded useful results. This was, in our opinion, the case for the very extensive retrospective studies after the Chernobyl accident, and for studies on populations exposed to chemicals where often the wrong tests were used for the wrong exposure and the wrong conclusions drawn by media have created much unnecessary anxiety."*

From Bender *et al.* (1988) (Based on a report of a Committee that was established by This paper is based on a report of a Committee that was established by a request of the National Cancer Institute to Oak Ridge Associated Universities' Medical and Health Sciences Division to evaluate devices and techniques that may be useful in determining and quantifying previous radiation exposures):

*"Large exposures of populations can also be detected even several decades after their exposure, but only in the case of populations, and of large doses (of the order of 100 to several hundred rad). The working group does not believe that cytogenetic measurements can detect internal doses from fallout radionuclides in individuals unless these are very large."*

*"The precision with which low doses can be detected in individuals, or even higher ones in individuals sampled long after their exposure, is less clear. Unfortunately, this is precisely the problem presented by the exposed veteran populations: their exposures occurred long ago and, from the physical evidence available (film badges, dose reconstructions, etc.), seem likely in most cases to fall in the low-dose category ( $\sim < 10$  rad). Thus the pertinent question is, with what precision can small exposures be detected in individuals whose lymphocytes are sampled decades afterward? Put another way, we may ask what confidence we should have in concluding that an individual was*

*indeed exposed several decades earlier if we observe some particular number of aberrations in a sample of a certain number of cells."*

*"Additional research on the problem is clearly needed, but at the moment it appears unlikely that determination of chromosomal aberration frequencies in peripheral blood lymphocytes will prove a useful method of determining ionizing radiation doses to individual veterans (though it might prove useful in showing that doses to veterans as a population were not greatly in excess of those presently estimated)."*

## References

- Bender, MA, Awa, AA, Brooks, AL, Evans, HJ, Groer, PG, Littlefield, LG, Pereira, C, Preston, RJ and Wachholz, BW (1988). Current status of cytogenetic procedures to detect and quantify previous exposures to radiation. *Mutat Res* 196:103-59.
- Hamasaki, K, Kusunoki, Y, Nakashima, E, Takahashi, N, Nakachi, K, Nakamura, N and Kodama, Y (2009). Clonally expanded T lymphocytes from atomic bomb survivors in vitro show no evidence of cytogenetic instability. *Radiat Res* 172:234-43.
- Hande, MP, Azizova, TV, Geard, CR, Burak, LE, Mitchell, CR, Khokhryakov, VF, Vasilenko, EK and Brenner, DJ (2003). Past exposure to densely ionizing radiation leaves a unique permanent signature in the genome. *Am J Hum Genet* 72:1162-70.
- Kleinerman, RA, Romanyukha, AA, Schauer, DA and Tucker, JD (2006). Retrospective assessment of radiation exposure using biological dosimetry: chromosome painting, electron paramagnetic resonance and the glycophorin a mutation assay. *Radiat Res* 166:287-302.
- Lazutka, JR, Lekevicius, R, Dedonyte, V, Maciuleviciute-Gervers, L, Mierauskiene, J, Rudaitiene, S and Slapsyte, G (1999). Chromosomal aberrations and sister-chromatid exchanges in Lithuanian populations: effects of occupational and environmental exposures. *Mutat Res* 445:225-39.
- Leonard, A, Rueff, J, Gerber, GB and Leonard, ED (2005). Usefulness and limits of biological dosimetry based on cytogenetic methods. *Radiat Prot Dosimetry* 115:448-54.
- McLean, AR and Michie, CA (1995). In vivo estimates of division and death rates of human T lymphocytes. *Proc Natl Acad Sci U S A* 92:3707-11.
- Muirhead, CR, Bingham, D, Haylock, RGE, O'Hagan, JA, Goodill, AA, Berridge, GLC, English, MA, Hunter, N and Kendall, GM (2003). Mortality and cancer incidence 1952-1998 in UK participants in the UK atmospheric nuclear weapons tests and experimental programmes. National Radiation Protection Board W27.
- Prabhavathi, PA, Fatima, SK, Padmavathi, P, Kumari, CK and Reddy, PP (1995). Sister-chromatid exchanges in nuclear fuel workers. *Mutat Res* 347:31-5.
- Simon, SL, Bouville, A and Kleinerman, R (2010). Current use and future needs of biodosimetry in studies of long-term health risk following radiation exposure. *Health Phys* 98:109-17.
- Whitehouse, CA, Edwards, AA, Tawn, EJ, Stephan, G, Oestreicher, U, Moquet, JE, Lloyd, DC, Roy, L, Voisin, P, Lindholm, C, Barquinero, J, Barrios, L, Caballin, MR, Darroudi, F and Fomina, J (2005). Translocation yields in peripheral blood lymphocytes from control populations. *Int J Radiat Biol* 81:139-45.

Ian Morison 14/7/2010  
Professor of Pathology  
Department of Pathology  
Dunedin School of Medicine  
University of Otago

**Qualifications**

B Med Sc (Otago)	1982
MB ChB (Otago)	1984
Medical Registration - the Medical Council of New Zealand	1985
Unrestricted Medical Licensure for Washington State	1989
American Board of Pathology Certificate in Clinical Pathology	1989
FRCPA (Haematology)	1992
Specialist Registration (Pathology) - the Medical Council of NZ	1993
PhD (Biochemistry, Otago)	1996

**Comment:**

I assessed these reports from the perspective of a Clinical Pathologist with approximately 20 years of experience in the use and assessment of diagnostic tests. During this time I have seen tests come and go with alarming frequency. Even tests that appear well validated and which we learn to trust can fail when tested on clinically appropriate patient and controls. During my professional experience I have, therefore gained a degree of distrust of new and incompletely characterised tests: this report reflects some of this mistrust. My report is based on my interpretation of the current literature on radiation biodosimetry. I have not discussed the original study, or my report, with any colleagues.



29 March 2010

Professor John Campbell  
Chairman – Ministerial Advisory Group on Veterans' Health  
PO Box 5146  
Wellington  
New Zealand  
Cc Mr Matt Scott  
Secretary

Dear Professor Campbell,

Thank you very much for letter on 24 February regarding the New Zealand Nuclear Test Veterans' Study. I am delighted to share with you my opinion on the two scientific reports written by the Institute of Molecular Biosciences at Massey University.

#### Research question

In both studies, the authors have used cytogenetic biomarkers to analyse the delayed effects of exposure to ionising radiation among a cohort of New Zealand Nuclear Test Veterans. This approach is straightforward and, in my opinion, represents the only way to evaluate the abovementioned effects. Given that the total number of exposed New Zealand Nuclear Test Veterans is relatively small (~1000, see Pearce *et al.*, 1990, *BMJ* **300**, 1161), the statistical power of any epidemiological study on this cohort is limited. On the other hand, the cytogenetic biomarkers can provide important information regarding the doses of exposure and also allow predicting the genetic consequences of exposure for this cohort.

#### Population sampling

The authors have collected blood samples from 50 male irradiated veterans and 50 non-irradiated male ex-servicemen. Judging from the reports, the participants were carefully selected and matched by age and other confounding factors. These two groups were used in the both studies. For the purpose of retrospective biodosimetry the sample size is appropriate.

#### Experimental techniques

In the first study (Rowland *et al.*, 2005, *New Zealand Nuclear Test Veterans' Study - a Pilot Project*), the authors used the sister chromatid exchange (SCE) assay. This is well-established cytogenetic technique which measures the incidence of *de novo* chromatid exchanges occurring in replicating cells. Given the high spontaneous frequency of SCE, this technique requires the scoring of a relatively small number of cells (~30-50 per individual, Albertini *et al.*, 2000, *Mutat Res* **463**, 111) to detect significant changes in their frequency in the exposed individuals. As on average ~30 and ~40 cells were respectively scored in the exposed and control subjects, the authors therefore followed the above-mentioned IPCS guidelines for the monitoring of genotoxic effects. I am slightly concerned with the fact that, on average, less cell per individual were analysed in the exposed group.

In the following study (Rowland *et al.*, 2007, *New Zealand Nuclear Test Veterans' Study - a Cytogenetic Analysis*), the authors used a number of well-established cytogenetic techniques, including the G2 assay,

micronucleus (MN) test and multicolour fluorescent *in situ* hybridisation (mFISH). I welcome the authors' choice of the techniques, as the G2 and MN tests detect the *de novo* cytogenetic alterations, whereas the mFISH approach mostly identifies chromosome translocations accumulated in the lymphocytes of exposed individuals over a considerable period of time. In other words, the authors attempted to establish whether the history of radiation exposure can be traced by the analysis of stable (mFISH) and unstable cytogenetic alterations. The number of cells analysed by the G2 test (~50 per exposed and control subjects) was sufficient to detect the differences in radiosensitivity between the two groups. In contrast, the number of binucleated lymphocytes scored per subject was on the borderline of the recommended range. Thus, the IPCS guidelines for the monitoring of genotoxic effects suggest scoring at least 1000-2000 cells per subject (Albertini *et al.*, 2000, *Mutat Res* 463, 111), whereas the mean number of cells scored by the authors was just ~1000 in the both groups. As far as the mFISH technique is concerned, again the mean number of cells scored per subject was below the internationally recognised standard. Thus, on average ~190 cells were caryotyped in controls and irradiated subjects, whereas the Panel of international experts suggested that at least 300 cells per subject should be typed (Sigurdson *et al.*, 2008, *Mutat Res* 652, 112).

## Results

According to the results of the first study, the frequency of SCE in the irradiated cohort significantly exceeded that in control. As the observed difference remained significant after an adjustment made for other potential confounding factors such as the number of cells scored per subject, smoking etc., and given that the two cohorts were matched by age, the authors' data clearly show that it is most probably attributed to radiation exposure. As the SCE technique, widely used for as a sensitive cytogenetic endpoint of testing the genotoxic risk, measures the incidence of *de novo* chromatid exchanges, the authors therefore concluded that the increased frequency of SCE may reflect the clastogenic activity of some long-lived radionuclides still remain in the body of veterans. Although this explanation may appear quite plausible, some other factors could also be implicated. For example, the delayed stimulation of SCE in the lymphocytes of exposed veterans is reminiscent of the phenomenon of radiation-induced genomic instability, where radiation can induce genetic changes in some of the descendants of a single irradiated cell, an effect that can persist for many cell generations (see Morgan, 2003, *Radiat Res* 159, 567). If correct, then the SCE data may provide an important evidence for the *in vivo* manifestation of radiation-induced genomic instability in humans.

In the following study the authors have provided strong evidence for the increased frequency of translocations measured by the mFISH technique in the cohort of New Zealand Nuclear Test Veterans. In contrast, the frequency of MN and the yield of DNA breaks measured by the G2 test in this group did not significantly differ from those in control. Given that the frequency of translocations remains elevated over a considerable period of time since radiation exposure, these results provide a strong evidence for the long-term effects of exposure to the veterans. Similarly to the previous study, all necessary steps were taken to exclude a majority of confounding factors such as age, smoking etc. Using the translocation data, the authors also attempted to reconstruct the doses of exposure to the cohort of veterans. Although the *in vitro* dose-response was robustly established, I am not entirely convinced by the authors' data on dose reconstruction. First of all, except some rare cases of high-dose exposure, the translocation test is seldom used for the purpose of individual bio-dosimetry. This is mainly attributed to the very high inter-individual variation in the frequency of translocation in control and irradiated individuals. Besides, as already mentioned, the mean number of cells scored per subject in the both cohorts was quite low, which should even more affect the robustness of individual bio-dosimetry. |

would therefore be more comfortable seeing in the text the mean estimates of doses received by the whole group of veterans.

In summary, the authors have conducted a comprehensive cytogenetic analysis of a large cohort of New Zealand Nuclear Test Veterans. The main results of their study provide a very strong experimental evidence for the long-term effects of radiation exposure to this cohort and therefore raise the important issue of health-related consequences for the veterans, including cancer. As far as the hereditary risk for this cohort is concerned, it should be noted that, given the small number of exposed servicemen, the detection of radiation-induced mutations among their offspring is highly problematic.

Yours sincerely,



Professor Yuri E Dubrova  
Department of Genetics  
University of Leicester  
Leicester LE1 7RH  
UK

Tel: +44 116 252 5654  
E-mail: yed2@le.ac.uk

**John D Dockerty**  
**Epidemiologist and Public Health Physician**  
P O Box 6489  
Dunedin North.

July 20, 2010

Professor John Campbell  
Chairman, Ministerial Advisory Group on Veterans' Health  
P O Box 5146  
Wellington.

Dear Professor Campbell

Thank you for asking me to assess Professor Dubrova's review of the NZ Nuclear Test Veterans' Study; in particular to ascertain the degree to which participant selection and adjustment for confounders may have influenced the findings.

To complete my assessment I have reviewed the two original reports by Rowland and colleagues<sup>1</sup> and the letter sent to you on 29 March by Professor Dubrova.

The 2005 and 2007 reports were both from the same study which compared people exposed to ionising radiation from nuclear tests 'veterans' with 'controls'. The exposed group consisted of veterans who were on ships in proximity to nuclear tests conducted in the mid-Pacific Ocean during 1957-8, as described in the reports. The controls were ex-servicemen not there during those tests (they were mainly army people, and some police). The choice of groups seems appropriate, although the details of how the controls were selected are a bit brief. Some information is given about the participation rate among veterans who were approached to be in the study. The authors report that 6 of 50 selected eligible veterans withdrew (response 88%). The response rate among eligible controls is not given. Individuals in the exposed group smoked more than the controls (for 27 versus 17 years on average).

In the 2005 report, Sister Chromatid Exchanges were assessed in the two groups. After adjustment for smoking, there remained a significant ( $p=0.03$ ) increase in the frequency of chromosomal translocations in the veterans, though the absolute difference between the two groups was very small and would relate to an uncertain material impact on the men's health.

In the 2007 report, assays were conducted for G2 (chromosomal sensitivity to damage from irradiation), micronucleus (MN - chromosomal sensitivity to damage from irradiation), and multicolour fluorescent *in situ* hybridisation (mFish, chromosomal 'painting' to detect aberrations).

The G2 and MN assays showed no significant difference between the two groups, ie no difference in the DNA repair competence of veterans versus controls. The mFish assays showed a highly significant difference in the frequency of translocations between veterans and controls (29 versus 10 per 1000 cells in the respective groups,  $p<0.0001$ ). Smoking did not account for this

---

<sup>1</sup>Rowland RE, Podd JV, Wahab M. New Zealand Test Veterans' Study - a pilot project (Sister Chromatid Exchange). Institute of Molecular BioSciences, School of Psychology, Massey University 2005.

<sup>2</sup>Rowland RE, Podd JV, Wahab MA, Nickless EM, Parmentier C, M'Kacher R. New Zealand Test Veterans' Study - a Cytogenetic Analysis. Institute of Molecular BioSciences, School of Psychology, Massey University 2007.

difference. Other potential confounders were not assessed, beyond matching factors which included age, area of residence, gender and to some extent occupation.

In summary, the authors found statistically significantly elevated frequencies of some chromosomal anomalies in their study of 50 exposed veterans versus 50 non-exposed controls, some five decades after the exposure to ionising radiation from nuclear tests.

Weaknesses of the study design include the small numbers of veterans and controls included, the use of 'surrogate' endpoints (SCEs and translocations) which might not relate to health or vital status of the men, and the inadequately described participation rate for the control group. People who were too ill (and maybe also a sizeable number of deaths) were excluded. If exposed veterans were more likely to be ill or to die as a consequence of their exposure (hence be excluded) then this could cause a bias which might underestimate the impact of irradiation.

Relating to confounding, I don't think that any obvious measured (eg smoking) or unmeasured confounders could lead us to dismiss the results of the mFish analysis in particular. The SCE analysis showed a small absolute difference, which may not be so robust in the context of the methods. With multiple comparisons the adjusted p value for that of 0.03 is not definitive. The mFish result is much clearer and might indicate long-term radiation damage from the tests.

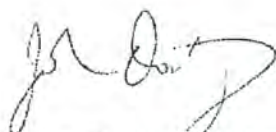
Some questions are not answered in this material: How were deaths *at any time since exposure* treated - there must have been some in the veterans' group since exposure? Should more work be done to follow up other endpoints (eg death and malignancies) among exposed veterans and suitably chosen controls? Could the work suggested by the authors in their 2007 report be conducted - ie another study looking at outcomes among the British and Fijian people who were also exposed as a consequence of the same nuclear tests?

Overall we have a small study with some careful design features and some other issues that make the results difficult to interpret or of uncertain consequence, as described above, especially in relation to what they might mean for the veterans' health in a tangible way.

Professor Dubrova gives a nice, thoughtful description of the research question and the methods and results of the genetic techniques used in the study. I assume from his address that his expertise is in genetics. When it comes to the epidemiological methods, I disagree with some of Prof Dubrova's conclusions (in his last paragraph). In particular: (1) this is not a 'large cohort' - but rather a study with a small sample size, as he points out earlier in his report (2) this does not give what I would regard as 'experimental' evidence, as it is an observational study and not a trial, (3) the evidence is not 'very strong' in my opinion, but it is suggestive of a difference which could even be an underestimate yet requires further confirmation, and (4) the consequences or implications for the men's health are not established by this type of research. While the findings are very interesting and warrant further investigation, on the basis of what has been reported we don't have enough evidence to attribute causality or to comment on the seriousness or implications of the findings. I have no conflict of interest.

With best wishes,

Yours sincerely,

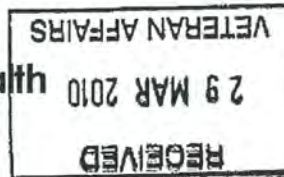


John D Docketty MB, ChB, Dip Comm Health, PhD, FAFPHM, FNZCPHM  
Epidemiologist and Public Health Physician



Australian Centre for  
Posttraumatic Mental Health

Mark Creamer, PhD  
Director  
Professor of Psychology  
Department of Psychiatry  
The University of Melbourne  
markcc@unimelb.edu.au



Level 1, 340 Albert Street  
East Melbourne  
Victoria Australia 3002

Tel: +61 3 9936 5100  
Fax: +61 3 9936 5199

acpmh-info@unimelb.edu.au  
www.acpmh.unimelb.edu.au

ABN: 47 720 569 707

Professor John Campbell,  
Chair, Ministerial Advisory Group on Veterans' Health,  
Allied Nationwide Finance Tower  
L 11, 142 Lambton Quay  
Wellington,  
New Zealand

March 20<sup>th</sup>, 2010

Dear Professor Campbell,

Thank you for your request to review the *New Zealand Nuclear Test Veterans Study: A Pilot Project (Psychological Impact)*. My colleague, Dr Ruth Parslow, and I have both reviewed the report in detail.

As you can see from the attached, we have concerns about the methodology and interpretation of the results in this report. Our concerns relate primarily to the representativeness of the samples, the lack of consideration given to factors other than the exposure, and the clinical relevance of the findings when compared to the broader NZ population.

I hope our review is useful to your Advisory Group.

Kind regards,

Mark Creamer.



Review of the *New Zealand Nuclear Test Veterans Study:*  
*A Pilot Project (Psychological Impact).*

## Executive Summary

The conclusions reached in the NZNTVA Report Pilot Project (Psychological Impact) rely strongly on having found statistically significant differences between 50 selected exposed veterans and 50 non-exposed controls. The report concludes that the sample of exposed veterans is in poorer mental health, and is experiencing more chronic health problems, than the non-exposed veterans. There are, however, several methodological flaws in the research and in the way in which the data have been interpreted. These raise significant questions regarding the conclusion that differences between these two groups are the result of having been exposed to nuclear testing during Operation Grapple in 1957-58 and the subsequent stress associated with this exposure.

Methodological issues of concern include:

- *The representativeness of the research samples:* There is sufficient evidence to question the representativeness of the veterans included in both samples; that is, whether the state of physical and mental health found in the two samples reliably reflects the state of the broader populations of interest. Further, there is evidence to suggest that the two groups may have differed on more than just their exposure to the nuclear tests, making it difficult to interpret any reported differences.
- *The lack of consideration given to other factors that could affect the health status of participants:* The research leaves reasonable doubt regarding the explanations for any reported health differences between the groups; if present, they are not necessarily explained by the nuclear test exposure. They may, for example, be due to pre-existing factors, health behaviours, sample selection, or a range of other variables not measured by this research or not incorporated into the appropriate analyses.

- *The failure to take into account how the groups rate on health measures when compared with older men in the wider New Zealand population:* Even if the findings of poorer mental and physical health in the exposed sample compared to the controls are not simply an artefact of the methodology, there is little evidence to suggest that they represent unusually poor health. The available data suggest that the mental and physical health status reported by the exposed group is comparable to the health status of other similarly aged New Zealand men drawn randomly from the general population.

Of course, these concerns should not be taken to mean that there have been no adverse effects of exposure to the nuclear tests, simply that this research does not prove the existence of those effects. It is disappointing to note that the report makes no reference at all to limitations of the study methodology or cautionary notes in interpreting the results. Instead, the results are presented without question or caveat, and the recommendations formulated accordingly.



## 1. Summary of study and report

### Participants

The report notes that an estimated 550 male veterans took part in Operation Grapple in 1957 and 1958, during which they were exposed to one or more nuclear bomb blasts. A subset of 200 of these was contacted by the NZNTVA researchers through this veterans organisation, of whom 151 expressed interest in participating in the study. The researchers identified a range of exclusion criteria which were applied to ensure participants' exposure to other potentially confounding risk factors comparable to Operation Grapple was minimised. These criteria included:

- Service in a theatre of war or nuclear-related area
- Exposure to toxic substances (including asbestos, oil/petrol fumes, road transport etc) for a year or more
- Having received radiation treatment or chemotherapy
- Aged over 75
- Air Force aircrew (potential exposure to cosmic radiation, other nuclear radiation exposure)
- Being too ill to participate

Of the 151 who expressed interest in participation, 88 met criteria for inclusion and a random sample of 50 of this group were selected as representing those who were exposed to nuclear testing in Operation Grapple.

For comparison, 135 individuals volunteered to be controls for this study. These people were recruited through a range of sources, including personal contacts. Exclusion criteria applying to the control group included those listed above, as well as the following additional criteria:

- Service in the NZ navy;
- Inability to match for age in a particular geographic region;
- Too high an education level;
- Recent immigration to New Zealand (as a means of controlling for background radiation exposure);
- Service that did not include compulsory military training.

With these criteria, a majority of potential participants (83) failed to meet inclusion requirements leaving only 52 able to participate. It was this low number of controls that led to the researchers reducing the number of exposed participants to 50.

A range of measures were then undertaken on the exposed and control groups including:

- the Mini-Mental State Examination (MMSE)
- the Geriatric Depression Scale (GDS)
- the Memory Assessment Clinics Self-Rating Scale (MAC-S)
- the SF-36 Health Survey.

As well, information on demographic attributes, chronic health problems and risk factors was collected from all participants.

#### **Findings:**

Preliminary analyses comparing the health measures of exposed and control groups identified significantly worse results for the exposed group on the following measures:

- depression (GDS score)
- a range of physical and mental health functioning measures (SF-36)
- distress about memory abilities (but not memory ability itself)

Participants' risk factors and demographic attributes were not analysed statistically in this report but included in the next set of analyses as potential confounding factors that should be controlled for. However, given the results they have provided, it is likely that exposed participants had significantly lower levels of education, less income and also had smoked at significantly higher levels than controls over their lifetime. Exposure to trauma scores calculated from the Trauma Exposure Scale (TES) for the exposed group were also higher, although the usefulness of this measure is questionable since '...there was some concern that some of the men may have misinterpreted the first item on the TES [Have you ever engaged in military combat?] and it was dropped from the analysis...'

When the analyses adjusted for these potential confounding variables, the exposed group were found to be more depressed, having significantly higher scores on the GDS ( $p < 0.001$ ) and poorer mental health scores as measured by the SF-36 ( $p = 0.04$ ). Exposed participants also reported significantly higher levels of cancer and chronic skin conditions compared with the control group.

Based on these findings, the authors recommend that those NZ veterans exposed to nuclear radiation as a result of Operation Grapple be appropriately assisted to help them cope with the chronic stress they are experiencing.

## **2. Assessment of analyses and conclusions**

This study is based on a relatively small sample of exposed veterans and non-exposed controls. On a small number of measures, the exposed group as a whole performs worse than the control group as a whole. There are three questions that these findings bring to the fore:

1. How representative are the exposed and control groups of the larger populations they are drawn from?
2. Can these differences be explained by factors that have not been accounted for in the study?
3. What do these differences mean for the physical and mental health of the exposed group, other than indicating that there is a difference between them and the control group that has a small likelihood of occurring only by chance?

---

### **How representative are the exposed and control groups of the larger populations from which they are drawn?**

The researchers sent invitations to participate to 200 of the 550 or so service men who participated in the nuclear tests. Obviously some would be deceased, but the number contacted is still less than 40% of the total exposed groups. This figure would be no problem if they were selected randomly. In this case, however, they were selected through a veterans' association – already a highly select sample. It is reasonable to assume that veterans who remain strongly aligned with an association

such as this are likely to have strong views about the role of that particular experience in their lives. The extent to which they represent the total possible sample is unclear but anecdotal evidence suggests that members of ex-service organisations tend to have poorer physical and mental health than the broader veteran population. A better design would have been a random sample drawn from nominal rolls of those involved in the nuclear tests.

The control group, on the other hand, was recruited through a range of sources including “personal contacts” and included not only people who had served in the Defence Forces, but also in the police or some other form of “compulsory military training”. The goal of a control group, of course, is to ensure that they are as alike as possible to the population of interest except in the one area of study focus (in this case, exposure to the nuclear tests). It is always difficult to recruit an appropriate control group, but there does seem to be danger in this sample that they are not necessarily comparable with the exposed group in areas that might constitute vulnerability for poor physical and mental health. In other words, they may well have been a healthier sample regardless of nuclear exposure. This hypothesis is supported by the comparisons with the general New Zealand population reported below. Unless we can be sure that the control group has comparable risk, we cannot be sure that the nuclear exposure is the explanation for any differences. A better design would have been a random sample drawn from nominal rolls.

The researchers have applied a large list of exclusion criteria in their attempts to select samples whose only exposure to nuclear risk was that of Operation Grapple. By using such broad exclusion criteria (for example, service in a theatre of war, having received radiation treatment or chemotherapy), they are left with a relatively small group who have had minimal risk of genetic damage outside their experiences in 1955. While this may not have a large effect on the representativeness of the exposed sample (whose total numbers are limited), it is likely to have a greater effect on the extent to which those in the control group can be said to adequately represent non-exposed military personnel. Those in the control group were also further restricted by additional exclusion criteria that were not applied to the exposed group. It is not surprising, then, that a higher percentage of potential controls were excluded (only 52 of 135 accepted).

In summary, there is sufficient evidence to question the representativeness of the veterans included in both samples; that is, whether the state of physical and mental health found in the two samples

reliably reflects the state of the broader populations of interest. Further, there is evidence to suggest that the two groups may have differed on more than just their exposure to the nuclear tests, making it difficult to interpret any reported differences. Thus, it is not possible to state definitively that veterans who were exposed have suffered adverse health effects as a result.

**Can these differences be explained by factors that have not been accounted for in the study?**

This research found that, in a small number of measures, exposed individuals have poorer health than non-exposed. The measures include the Geriatric Depression Scale, the SF-36, and prevalence of cancer and of chronic skin conditions.

*Mental Health*

The first two of these measures indicate poorer mental health. It is well known that mental health is negatively affected by a range of factors including personality attributes (e.g., higher levels of neuroticism), biological factors (e.g., vascular factors), and use of some medications (see Vink et al, *Risk factors for anxiety and depression in the elderly: A review, J Affect Disorders*, 2008, 106:29-44). These factors were not measured and could help to explain the poorer mental health of those in the exposed group. In large randomly selected samples, we would argue that these differences should be equal across both groups. These samples, however, were small and not randomly selected. Attributing the differences in mental health status of the exposed group only to their exposure and consequent concerns ignores other potential risk factors affecting mental health of these groups.

*Cancer and chronic skin conditions:*

One of the criteria by which potential participants were excluded was “having received radiation treatment or chemotherapy” (p11) – the common cancer treatments. The logic of this exclusion is not entirely clear, since it limits the type and severity of cancers expected to be reported by participants entering this study. It is distinctly possible that one group (perhaps the exposed group) would have shown higher cancer rates if anyone who had received treatment for cancer had not been excluded.

The researchers have already reported that the exposed sample in the study has a much higher level of smoking than the control group. Tobacco smoking is a risk factor for multiple chronic skin conditions (see for example, Van der Straten, Carrasco et al, Tobacco use and skin disease, *Southern Medical Journal* 2001; 94:621-34). Other research has also identified tobacco smoking as a risk factor for multiple skin problems including squamous cell carcinoma (Freiman, Bird et al, Cutaneous effects of smoking, *Journal of Cutaneous Medicine and Surgery*, 2004; 8:415-23.) While the researchers controlled for some potential confounding factors including tobacco smoking when they compared participants' scores for depression, mental and physical health and memory, they have only presented simple unadjusted analyses when comparing prevalence of chronic health problems in the two groups. Their findings that exposed participants had significantly poorer health in two skin-related diseases in a list of 23 chronic health conditions could be attributable, at least in part, to the significant differences in risky health behaviour reported by the two groups. Of course, it is possible that the increased smoking was a result of their exposure (or other demographic or military factors) but the report does not address this issue.

In summary, the research leaves reasonable doubt regarding the explanations for any reported health differences between the groups; if present, they are not necessarily explained by the nuclear test exposure. They may, for example, be due to pre-existing factors, health behaviours, sample selection, or a range of other variables not measured by this research.

**What do these differences mean for the physical and mental health of the exposed group other than indicating that there a difference between them and the control group?**

In this final section, the findings are examined not by identifying statistical differences between these two small groups, but by exploring what these findings may indicate about the well-being of the exposed and control groups when compared with population norms. A small statistical difference between these two groups is of interest from a research perspective but may have little relevance for the well-being of either group if their overall measures fit within the expected range of such measures found in the general population.

The exposed group had significantly higher levels of depressive symptoms compared with the control group. The unadjusted average score for exposed on the 15 item GDS was 3.92 with a standard error of 0.49. Without further details concerning the distribution of scores for this group, it could be expected that 95% of participants score in the range 2.96 to 4.88 on this scale. Does this represent an ill group of individuals? The GDS-15 has been tested in psychometric studies which have reported that individuals meeting diagnostic criteria for depression can be expected to score greater than 10, those with mild depression scoring between 5 and 10 and those scoring less than 5 come in the normal range (Sheil & Yesavage, Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version, *Clinical Gerontologist* 1986; 5:165-173. That is, 95% of exposed participants could be expected to score in the normal range on this measure. This would suggest that, at most, 1 or 2 individuals of the survey might screen positive for mild depression. It should be noted that in a recent New Zealand study of 208 older individuals drawn from the community, the mean GDS-15 score was 2.85 with standard error 0.19. Around 18% of this large sample recorded scores of between 5 and 10, and would probably have been experiencing mild depression (Andrew & Dulin, The relationship between self-reported health and mental health problems in older adults in New Zealand, *Aging and Mental Health*, 2007, 11:596-603)

The exposed group also reported significantly poorer mental health on the Mental Health Subscale (MHS) of the SF-36. The average unadjusted score for the MHS recorded by the 50 exposed individuals was 77.84 (SE 2.98). This measure has also been applied to a large sample of the New Zealand population which included 338 men aged between 65 and 74 years (Scott et al, SF-36 health survey reliability, validity and norms for New Zealand, *ANZ Journal of Public Health* 1999, 23:401-406). The mean MHS score for this age-sex group in this New Zealand study was calculated to be 83.8 (SE = 0.9). Using the same statistical testing as that applied in the NZNTVA report, there is no statistically significant difference in the MHS measures collected from the 50 participants in the exposed group of the NZNTVA study and those recorded for this larger group of older New Zealand men drawn from the civilian population ( $t=1.90$ ,  $p\sim 0.06$ ).

It has been noted previously in this review that selection process for the controls was more restricted than for the exposed subgroup, resulting in a greater number of exclusions and possibly in a healthier than usual subgroup of individuals. This is supported by comparing the mean MHS score

of controls with the general NZ population. The mean, unadjusted MHS score for the 50 controls was 90.24 with standard error of 1.48. The same statistical testing used previously indicates that these two groups, the community group from the New Zealand population and the 50 participants selected to be controls, are statistically significantly different with the control group having substantially better mental health than that reported by the equivalent age-sex group in the general population.

In summary, even if the findings of poorer mental and physical health in the exposed sample are not simply an artefact of the methodology, there is little evidence to suggest that they represent poor health states when compared to the broader population of age matched males in New Zealand.



## **Review of the study “New Zealand Nuclear Test Veterans’ Study – a pilot project (sister chromatid exchange)”**

**Drs Al Rowland, John Podd, and Mohammed Wahab**

### **A report presented to the New Zealand War Pensions Medical Research Trust Board, 2005**

This study describes a case-cohort study of New Zealand War Veterans who were involved in nuclear test detonations in the Pacific during the 1950s. The study seeks to examine evidence for chromosomal damage in these individuals compared to a matched cohort of similarly aged individuals who were not in the vicinity of the test sites at the time of the detonations. The measure that is chosen to gauge whether or not chromosomal damage has been sustained is a laboratory measure of sister chromatid exchange. This is a naturally occurring process that occurs at a basal level in all humans that becomes elevated in conditions in which chromosome breakage is increased.

The authors should be commended for their thorough and thoughtful approach to what is a difficult problem – measuring a biological effect to an exposure which was sustained in the distant past. I think the matching of the exposed and control groups was very good, although in retrospect one would have controlled for cigarette smoke exposure which is a potentially confounding variable.

The scientific design and methodology employed to measure sister chromatid exchange is excellent and I do not believe can be faulted. The results of the study indicate that the exposed and control groups exhibit a significantly different degree of sister chromatid exchange, with the exposed group demonstrating more elevated levels. The authors reasonably conclude that this is indicative of statistically significant elevations in chromosomal breakage, a known consequence of exposure to ionizing radiation.

The authors have endeavoured to account for their findings as artefactual in an honest and genuine fashion but failed to locate a confounding factor that in their mind explains their experimental findings. One such factor (cigarette smoking) is discounted since previous studies have shown that its effect on elevating SCE is not measurable a comparatively short time after smoking cessation. At another point in this report, however, the authors quote evidence that indicates that elevated SCE induced by exposure to ionizing radiation can persist for decades, not only because the radionuclides persist in the body, but also because affected blood lymphocytes

can have pronounced longevity. Although definitive proof is lacking either way this reviewer is left wondering if the smoking has been definitively discounted as a confounding factor.

This observation aside, if the elevated level of SCE is presumed to be associated with radiation exposure in excess of 40 years prior to study, then what would the implications be? The first thing to note is that although the distribution of SCE is statistically different between the two groups studied (and this is examined in a reasonably rigorous statistical fashion), the magnitude of elevation of rate of SCE in the exposed group is still very modest. To my knowledge there is little in the literature that ascribes the size of this change to potential health outcomes. This would, therefore, imply that although a biological difference between these two groups has been detected, its magnitude may be of insufficient size to have any untoward affect on the health of the exposed group.

At one point in this report the authors also suggest that the offspring of the exposed group be monitored on account of their father's radiation exposure. For the same reasons above there is no evidence to suggest that this magnitude of change will impact adversely upon the health of the offspring of the exposed group. It must be conceded however that what is being measured here is the vestige of a long-gone event. It is entirely plausible that the degree of SCE may have been distinctly abnormal immediately after the exposure.

In summary, I think that these authors have certainly demonstrated a biological difference between these two study groups. There is no way of knowing whether this difference would have been far greater and clearer if it had been studied at a time nearer to the exposure in the late 1950s. Confounders may explain this difference but they are hard to identify. I am unconvinced that cigarette smoking has been formerly excluded as a confounder in this study, although I must concur with the authors that the weight of evidence lies against it being the explanation for the effect observed.

There are other genetic measures of genomic damage that have been used in the wake of other radiation releases into the atmosphere, most notably micro-satellite instability. I would think that a case could be made to study these cohorts further using another technique such as this to further support their findings if a causal link was to be strengthened.

Although cautious conservative practice may advocate increased medical monitoring of the exposed veterans, there is no evidence that any adverse effect of this magnitude is conferred transgenerationally and, therefore, I think any recommendation for enhanced medical surveillance of the offspring of the exposed group is not sustainable at this time.

**Stephen Robertson**  
**Professor of Paediatric Genetics**

**8 December 2005**