Technical Data

Populations Living near the Semipalatinsk Test Site: A New Cohort for Assessing Health Effects from Protracted Exposures to Low-to-moderate Doses of Radiation

Shinji YOSHINAGA,*1,†,‡ Jun'ichi ISHIDA,*2 Jiro INABA,*2 Zhanat BAIGAZINOV,*3 Marina V. BEREZINA,*3,††
Gulmara T. KENZHINA,*3 Sergei A. BEREZIN*3 and Toshiaki OGIU*2,††

(Received on July 3, 2018)

(Accepted on August 6, 2018)

Although numerous epidemiological studies have demonstrated an increased risk of cancer and non-cancer diseases associated with acute exposures to high doses of radiation, possible effects of protracted exposures to low-to-moderate doses of radiation remain to be investigated. Several epidemiological studies have been conducted for residents living near the Semipalatinsk Test Site in the former Soviet Union who were exposed to protracted low-to-moderate doses of both external and internal radiations. However, small population size and other limitations of the previous studies preclude a precise evaluation of the health effects. An international joint study of Kazakhstan and Japan was initiated in 2001 to investigate the health effects among residents of the exposed areas based on a new large-scale cohort of the Semipalatinsk populations. In the joint study, information was obtained on the name, date of birth, sex, residential history and vital status of individuals living in the former Semipalatinsk oblast and other exposed areas using archived and official data. By March 2009, data of 131,723 residents were collected, among which the vital status of 41,128 residents (30.7%) was ascertained. The completeness of the ascertainment depended on the place of residence with lower completeness for control areas mainly because of emigration. Although any analysis based on retrospective follow-up data is limited because of potential selection bias, the sub-population of our established cohort will be a useful resource for future prospective studies.

KEY WORDS: epidemiological study, Semipalatinsk Test Site, protracted exposure, cancer risk, health effect, cohort, mortality.

I INTRODUCTION

Although numerous epidemiological studies have demonstrated an increased risk of cancer and non-cancer diseases associated with acute exposures to relatively high doses of ionizing radiation,^{1, 2)} possible effects of protracted exposures to low-to-moderate doses of radiation remain to be investigated. Very few populations worldwide can allow the evaluation of the health effects of such radiation exposures, which is a considerable social and scientific concern.

The United Nations Scientific Committee on the Effects of Atomic Radiation reviewed currently available epidemiological findings on radiation and cancer in its 2006 report.¹⁾ The report included the results of new or updated studies on atomic bomb survivors and other nine populations, including residents living near the Semipalatinsk Test Site (STS) of the former Soviet Union. In those areas, 500,000–1,000,000 residents of all ages and both sexes were exposed to external and internal radiations resulting from radioactive fallout over an extended period of time. A number of epidemiological studies have been conducted in these areas, which were based on a historical cohort originally established by the local research institution “Dispensary No. 4,” which was later succeeded by the Research Institute for Radiation Medicine and Ecology, Kazakhstan. The members of the historical cohort comprising approximately 20,000 residents (9,850 in 10 exposed areas and 9,604 in 6 control areas) who were born before July 1961 and had permanently lived in one of the study areas were retrospectively followed up for cancer and non-cancer mortality.^{3, 4)} BAUER et al. showed a significant dose-related increase in solid cancer risk mortality,⁵⁾ whereas GROSCHKE et al. did not find any detectable risk of radiation-related mortality from cardiovascular diseases in this cohort.⁶⁾ However, the relatively small population size and other

*1 Center for Radiation Protection Knowledge, National Institute of Radiological Sciences, National Institutes for Quantum and Radiological Science and Technology; 4-9-1, Anagawa, Inage-ku, Chiba-shi, Chiba 263-8555, Japan.

*2 Radiation Effects Association; 1-9-16, Kajicho, Chiyoda-ku, Tokyo 101-0044, Japan.

*3 National Nuclear Center; 2 Krasnoarmayskaya, Kurchatov, 071100, Republic of Kazakhstan.

† Present Affiliation: Research Institute for Radiation Biology and Medicine, Hiroshima University; 1-2-3 Kasumi, Minami-ku, Hiroshima-shi, Hiroshima 734-8553, Japan.

†† They passed away during the writing of this paper. They remain authors as a reflection of their substantial and valued contribution.

Corresponding author; E-mail: syoshinaga@hiroshima-u.ac.jp

study limitations inherent to retrospective follow-up studies precluded a precise evaluation of the health effects of radiation exposure among these individuals.

In 2001, the Radiation Effects Association (REA) of Japan initiated the “Study on Health Effects of Radiation on Residents near the Former Semipalatinsk Nuclear Test Site” in collaboration with the Center for Study and Protection from Radiation Effects of Kazakhstan. Thus, a much larger cohort of the Semipalatinsk population was subsequently established to conduct epidemiological studies. In 2004, the National Nuclear Centre (NNC) took over the Center for Study and Protection from Radiation Effects and since then REA and NNC jointly performed the study. In this paper, we describe cohort profiles as well as methods to obtain relevant information for a future prospective study.

II METHODS

1. Study areas and study population

Since World War II, more than 700 nuclear tests have been conducted in the former Soviet Union, including the STS, covering an area of 18,500 km². A total of 456 nuclear tests were conducted at the STS between 1949 and 1989, including 86 tests in the air and 30 tests on the ground.⁷⁾ Thirteen of the underground tests are believed to have caused a substantial release of radioactive nuclides into the atmosphere, while five of the surface tests were not successful and resulted in the substantial contaminations of radioactive materials into the environment.⁸⁾ These releases, together with releases from the tests performed between 1949 and 1962, caused radioactive plumes by which residents were exposed to considerable amounts of radiation, particularly in the former Semipalatinsk oblast. In this study, we focused on the following areas with more or fewer contamination due to fallout from these nuclear tests; 1) Semipalatinsk I: four of the heavily exposed villages, including Dolon, Znamenka, Karaul and Kainar, in the former Semipalatinsk oblast, 2) Semipalatinsk II: other 11 exposed villages in the southern part of the Beskaragai Region in the former Semipalatinsk oblast, including Mostik, Cheremushki, Budene, Bolshaya Vladimirovka, Malaya Vladimirovka, Izvestka, Besterek, Kanonerka, Semiyarka, Semenovka and Sosnovka, 3) Pavlodar I: three of the less exposed villages located in the centre of Pavlodar oblast, including Kenzhekol, Kenes and Zhanaaul, and 4) Pavlodar II: three other less exposed villages located in northern and eastern areas of Pavlodar oblast, including Kachiry, Irtysh and Sharbakty. Individuals who were confirmed to have ever lived in these areas between 1949 and 1963 were identified and geographed in our database as part of the study population. The geographical locations of the study areas are shown in Fig. 1.

2. Data sources

Primary information sources for establishing the study cohort were two sets of official data, one from local archives and the other from citizen registration offices called ZAGS. First, we made a roster for individuals living in the study areas by reviewing archived data such as household books of residents, which were updated every 3 years and kept in local

Fig. 1 Geographical location of study areas near the Semipalatinsk Test Site (STS).

Note: Red and blue circles indicate exposed villages and control villages, respectively.

administrations. Such information included name, date of birth, sex, address, ethnicity, family structure, occupation, date and place of relocation from the previous area of residence and year of death, when applicable. However, date of birth and other information were sometimes lacking in the archived data. Additionally, updated information was no longer available at the original local administration if the residents married or changed residence location to a different district.

Citizen registration offices collected other primary information that was useful to establishing the cohort as well as for ascertaining the residential history and vital status of the individuals. The information included name, address, date of birth, parents' names, date of birth, date of death, physician-certified cause of death, name of spouse, date of marriage and former name. Updated information such as new address was no longer available at the original citizen registration offices after residents changed residence location to a different district. We verified and complemented the data collected from local archives by checking the data at citizen registration offices.

It should be noted that the procedures for establishing the study cohort were slightly different for the four population groups described above. For the Semipalatinsk I and Pavlodar I populations, a roster of residents for the period between 1949 and 1963 was first prepared based on the archived data of household books, and additional data from household books up to 2000 were supplementarily used. For the Semipalatinsk II and Pavlodar II populations, however, a roster of residents

during the period between 1961 and 1963 was first prepared on the basis of records at citizen registration offices, and additional data were obtained from household books recorded between 1949 and 1963.

3. Ascertainment of residential history and vital status

Residential history for individuals was constructed by compiling registration records from the citizen registration office. A total of 131,723 individuals were confirmed to have lived in the study areas for shorter or longer periods between 1949 and 1963, and their records were registered into our database. Through periodical visits to citizen registration offices, the vital status of these individuals was updated until 31 March 2009. For a limited number of subjects, interviews with living subjects, military records and medical records were also used to supplement the information on the residential history or vital status. Information on vital status and causes of death was no longer available at the original citizen registration offices after residents changed residence location to a different district. Thus, catchment areas for mortality follow-up were limited to the study areas. For deceased individuals, information on cause of death was obtained from citizen registration offices and was coded using the International Classification of Disease (ICD), version 10.⁹⁾

III RESULTS

1. Demographic characteristics

We classified the individuals into four groups according to the villages in which they resided for the longest period

between 1949 and 1963: Semipalatinsk I, Semipalatinsk II, Pavlodar I and Pavlodar II. As described earlier, the first two groups refer to exposed populations, while the last two groups refer to control populations. Out of 131,723 individuals, we had records for 9,431 individuals living in one of the 15 villages in Semipalatinsk I or II, but the main residence could not be established as either one of the 15 villages because information on residential history between 1949 and 1963 was substantially limited for them. For this reason, we excluded these 9,431 individuals from further analyses. Thus, the resulting study population was of 122,292 individuals whose data were used for descriptive analysis.

Table 1 shows the number of individuals from each population included in these four groups by sex, year of birth and ethnicity. There were slightly more females ($n = 64,009$, 52.3%) than males ($n = 58,283$, 47.7%), and the sex ratio was almost constant in the four population groups. For the total population, the majority of the subjects were Russians and Kazakhs (37.9% and 32.2%), followed by Ukrainians (12.1%). Kazakhs were more predominant (66.9%) in the Semipalatinsk I group.

2. Vital status and cause of death

Vital status (whether subjects were dead or alive) was ascertained by data obtained from public archives and citizen registration offices as described above. **Table 2** shows the vital status of the study population as of 31 March 2009. A total of 47,537 subjects (30.7%) were considered to be either

Table 1 Demographic characteristics of the study population.

	Semipalatinsk I	Semipalatinsk II	Pavlodar I	Pavlodar II	Total
Sex					
Male	12,410	12,218	10,662	22,993	58,283 (47.7%)
Female	13,095	14,212	11,393	25,309	64,009 (52.3%)
Year of birth					
–1919	5,340	6,212	4,633	9,950	26,135 (21.4%)
1920–1929	2,962	3,560	2,706	6,782	16,010 (13.1%)
1930–1939	4,691	5,363	4,423	10,079	24,556 (20.1%)
1940–1949	4,213	3,994	3,381	7,344	18,932 (15.5%)
1950–1959	5,698	5,581	5,180	11,342	28,071 (23.0%)
1960–1963	2,302	1,661	1,708	2,744	8,415 (6.9%)
unknown	29	59	24	61	173 (0.1%)
Ethnicity					
Kazakh	17,057	9,374	6,537	6,352	39,320 (32.2%)
Russian	4,670	12,294	7,320	22,075	46,359 (37.9%)
Ukraine	890	817	2,751	10,341	14,799 (12.1%)
German	761	2,269	2,297	4,987	10,314 (8.4%)
Others	2,127	1,676	3,150	4,547	11,500 (9.4%)
Total	25,505	26,430	22,055	48,302	122,292 (100%)

Table 2 Vital status of the study population as of March 31, 2009.

	Semipalatinsk I	Semipalatinsk II	Pavlodar I	Pavlodar II	Total
Alive	5,871	2,125	1,936	3,613	13,545 (11.1%)
Dead	7,006	4,962	3,755	8,266	23,989 (19.6%)
Lost to follow-up	12,628	19,343	16,364	36,423	84,758 (69.3%)
Total	25,505	26,430	22,055	48,302	122,292 (100%)

alive ($n = 13,545$, 11.1%) or dead ($n = 23,989$, $n = 19.6\%$) while vital status was unknown before 1 January 2000 for 84,758 (69.3%) subjects mainly because they changed residence location or because of emigration. We treated them as “lost to follow-up” in the present study. There was a strong variation in completeness of vital status ascertainment by the four population groups (50.5%, Semipalatinsk I; 26.2%, Semipalatinsk II; 25.8%, Pavlodar I; 24.6%, Pavlodar II).

Causes of death were ascertained through the records from citizen registration offices for 23,471 subjects (97.8%) out of the 23,989 subjects who were recorded as deceased. Circulatory diseases and cancer were the leading causes of death as shown in **Table 3**. Roughly 40% of the total deaths were attributable to circulatory diseases, compared with 17% that were attributable to cancer. The proportion of cancer deaths was slightly higher among Semipalatinsk I and II populations (17.8% and 19.5%, respectively) than among Pavlodar I and II populations (12.2% and 16.6%, respectively).

IV DISCUSSIONS

By using official documents, we identified the 131,723 residents living near the STS, compared with roughly 20,000 residents who were retrospectively followed for cancer and non-cancer mortality in the previous studies of the historical cohort. The large population size is an advantage for an epidemiological study especially as we aimed to evaluate the effects of low-dose radiation. For this reason, we expanded the study area to 15 villages in our present survey from only 7 villages (Dolon, Znamenka, Karaul, Kainar, Mostik, Cheremushki and Kanonerka) included in the historical cohort. However, comparisons of both cohorts should be interpreted with caution regarding the following points.

First, the eligibility of the study population was slightly different between the two cohorts. In the new cohort, subjects had to live in the study areas at any given time between

January 1949 and December 1963. In contrast, in the historical cohort, subjects had to continue residing in the study areas until the end of 1962. Thus, the residence of the study subjects was less likely to be stable in the new cohort than in the historical cohort. Indeed, vital status was ascertained only for 41,128 residents (30.7%) in the present survey, and completeness of ascertainment was dependent on the location of residence, with lower completeness in Semipalatinsk II and Pavlodar I and II populations. This low completeness was generally related to individuals changing residence to other areas. However, under the current citizen registration system, it is almost impossible to trace residents who moved away from the study area. In this study, it is reasonable to think that subjects who died in the study areas were more likely to be identified as a cohort member through the retrospective procedures than living subjects who moved outside the study areas. Therefore, any analysis based on retrospective follow-up data is limited given the potential bias attributable to the subject selection process.

Second, accuracy of causes of death is a matter of concern for mortality study. We conducted a pathological review for 655 cancer cases which were registered in two cancer centers in the study areas (data not shown). The pathological review suggested that diagnoses were almost perfect for cancer cases, but uncertainty in causes of death remains for non-cancer diseases including circulatory diseases. In this study, causes of death were ascertained by official documents at the civil registration offices or archived records both of which are basic information sources of the cause-of-death register in the Semipalatinsk region.¹⁰⁾ Thus, the quality of information on causes of death in this study is comparable with that of the historical cohort study while underlying causes of death were coded according to ICD-9 for the historical cohort and to ICD-10 for the new cohort. The number of deaths from solid cancer and circulatory diseases were reported to be 889 (11.9%) and

Table 3 Distribution of selected causes of death among the study population.

	Semipalatinsk I	Semipalatinsk II	Pavlodar I	Pavlodar II	Total
All causes	6,679	4,774	3,753	8,265	23,471
Cancer	1,191	929	458	1,368	3,946
All solid	1,163	908	453	1,344	3,868
Esophagus	323	111	71	87	592
Stomach	229	173	125	363	890
Colon	17	18	7	23	65
Liver	112	71	29	112	324
Lung	158	192	71	258	679
Breast	25	43	12	36	116
Thyroid	5	3	3	5	16
HLP cancer*	35	30	10	46	121
Leukemia	28	21	5	24	78
Circulatory	2,486	2,487	1,077	3,552	9,602
IHD [#]	1,069	1,303	182	1,093	3,647
Cerebrovascular	593	524	179	758	2,054
Digestive	166	129	86	241	622
External causes	132	203	172	458	965

* Hematopoietic and lymphatic cancer

[#] Ischemic heart disease

3,211 (42.8%) out of 7,511 total deaths between 1969 and 1999 in the historical cohort.^{5,6} In contrast, the corresponding numbers were 3,868 (16.5%) and 9,602 (40.9%) out of 23,471 total deaths in the present study. The most frequent solid cancer sites in the new cohort were stomach (23% of all solid cancers), lung (18%) and oesophagus (15%) compared with oesophagus (36%), stomach (17%) and lung (15%) in the historical cohort. Differences between the new and historical cohorts regarding age, lifestyle and other factors seemed to result in differences in the causes of death.

Third, this study involved the calculation of individual dose estimates. The radiation dose received by the residents is a key issue in studying the health effects of radiation exposure. Many dosimetric works have been performed on the historical cohort,¹¹ and the mean cumulative external dose for residents in exposed areas was estimated to be 0.09 Gy by dose reconstruction models.⁶ We used a dose reconstruction model similar to that used in the historical cohort, which was originally developed by GORDEEV.¹² However, in the present study, some parameters were modified. Although the individual dose estimation has not been finalized yet in present study, preliminary calculations showed the residents in Semipalatinsk I and II received doses of comparable levels to the exposed areas evaluated in the historical cohort (data not shown).

Despite the limitations of the present study, the data of our established cohort, at least that of the sub-populations, will be useful resources for a future prospective study without selection bias related to identifying subjects. We believe that almost all the members of historical cohort living in the study areas of current study were included in the new cohort because the data sources for establishing cohorts were common for the historical and new cohorts. The international joint project "SEMI-NUC" was initiated in 2013 to assess the feasibility of a future prospective study using information from both the historical and new cohorts.¹³ A future study based on both datasets of the residents living near the STS will provide a unique opportunity for a better understanding of the health effects related to protracted exposures to low-to-moderate doses of radiation. The foundation of such a prospective study has been carefully established by the present work, and the construction of a highly reliable epidemiological database will prove to be an invaluable scientific tool for determining the health effects of low-to-moderate doses of radiation.

V CONCLUSIONS

Although the vital status data could only be obtained for 30% of the residents of the investigated areas, our established cohort has high potential to address questions on possible health effects associated with protracted exposures to low-to-moderate doses of radiation. A future epidemiological study of the unique Semipalatinsk cohort will provide insight for a better understanding and quantification of radiation risks. Our database will be a useful tool for future prospective studies on the health effects of radiation in these populations.

ACKNOWLEDGEMENT

This study was sponsored by the Japanese Ministry of Education, Culture, Sports, Science and Technology. We are grateful to the members of the Semipalatinsk Health Effects Research Committee at the Radiation Effect Association for their discussions and recommendations for this study. We also thank the members of the Center for Study and Protection from Radiation Effects and the National Nuclear Center of the Republic of Kazakhstan for their efforts in data collection. All authors have no conflict of interest to be declared.

REFERENCES

- 1) United Nations Scientific Committee on the Effects of Atomic Radiation; Effects of ionizing radiation (UNSCEAR 2006 Report Volume 1), United Nations, New York, Available at: http://www.unscear.org/docs/reports/2006/07-82087_Report_2006_Web.pdf, Accessed 21 June 2018.
- 2) K. OZASA, Y. SHIMIZU, A. SUYAMA, F. KASAGI, M. SODA, et al.; Studies of the mortality of atomic bomb survivors, Report 14, 1950–2003: an overview of cancer and noncancer diseases. *Radiat. Res.*, **177**, 229–243 (2012).
- 3) B. I. GUSEV, Z. N. ABYLKASSIMOVA and K. N. APSALIKOV; The Semipalatinsk nuclear test site: a first assessment of the radiological situation and the test-related radiation doses in the surrounding territories, *Radiat. Environ. Biophys.*, **36**, 201–204 (1997).
- 4) B. GROSCHKE, C. LAND, S. BAUER, L. M. PIVINA, Z. N. ABYLKASSIMOVA, et al.; Fallout from nuclear tests: health effects in Kazakhstan, *Radiat. Environ. Biophys.*, **41**, 75–80 (2002).
- 5) S. BAUER, B. I. GUSEV, L. M. PIVINA, K. N. APSALIKOV and B. GROSCHKE; Radiation exposure due to local fallout from Soviet atmospheric nuclear weapons testing in Kazakhstan: solid cancer mortality in the Semipalatinsk historical cohort, 1960–1999, *Radiat. Res.*, **164**, 409–419 (2009).
- 6) B. GROSCHKE, D. T. LACKLAND, C. E. LAND, S. L. SIMON, K. N. APSALIKOV, et al.; Mortality from cardiovascular diseases in the Semipalatinsk historical cohort, 1960–1999, and its relationship to radiation exposure, *Radiat. Res.*, **176**, 660–669 (2011).
- 7) United Nations Scientific Committee on the Effects of Atomic Radiation; Sources and effects of ionizing radiation (UNSCEAR 2008 Report Volume 1), United Nations, New York, Available at: http://www.unscear.org/docs/reports/2008/09-86753_Report_2008_Annex_B.pdf, Accessed 21 June 2018.
- 8) V. A. LOGACHEV, N. P. VOLOSHIN, A. M. MATUSHCHENKO, L. A. MIKHAIKINA, S. G. SMAGULOV, et al.; "Nuclear tests of the USSR. Semipalatinsk test site: ensuring the general and radiation safety of nuclear tests", p.330 (1997), Moscow, Medbioextrem.
- 9) World Health Organization; International statistical classification of diseases and related health problems, 10th revision, Volume 2, Geneva (1994).
- 10) R. A. WINKELMANN, F. D. TRETIAKOV, N. V. STARTSEV, I. B. KOLYADO, B. I. GUSEV, et al.; Cause-of-death registers

in radiation-contaminated areas of the Russian Federation and Kazakhstan, *Radiat. Environ. Biophys.*, **41** 61–68 (2002).

- 11) V. F. STEPANENKO, M. HOSHI, I. K. BAILIFF, I. B. KOLYADO, B. I. GUSEV, et al.; Around Semipalatinsk nuclear test site: progress of dose estimations relevant to the consequences of nuclear tests, *J. Radiat. Res.*, **47**, Suppl A1–A13 (2006).
- 12) K. I. GORDEEV, et al.; Evaluation of integral absorbed and effective doses for population living permanently at

the radioactive traces resulted from atmospheric nuclear explosions: Methodological Handbook, Ministry of Health of Russian Federation (2001).

- 13) SEMI-NUC project: prospective cohort study of residents near the Semipalatinsk test site —feasibility assessment, International Agency for Research on Cancer website, Available at: <http://semi-nuc.iarc.fr>, Accessed 21 June 2018.