

Half-century archives of occupational medical data on French nuclear workers: a dusty warehouse or gold mine for epidemiological research?

Jerome-Philippe Garsi¹, Eric Samson¹, Laetitia Chablais¹, Sergey Zhivin¹, Christine Niogret², Dominique Laurier¹, and Irina Guseva Canu¹

Laboratoire d'épidémiologie, Institut de Radioprotection et de Sécurité Nucléaire, Fontenay-aux-Roses¹, AREVA NC, Occupational Medicine Department, Pierrelatte², France

Received in November 2013

CrossChecked in November 2013

Accepted in November 2014

This article discusses the availability and completeness of medical data on workers from the AREVA NC Pierrelatte nuclear plant and their possible use in epidemiological research on cardiovascular and metabolic disorders related to internal exposure to uranium. We created a computer database from files on 394 eligible workers included in an ongoing nested case-control study from a larger cohort of 2897 French nuclear workers. For each worker, we collected records of previous employment, job positions, job descriptions, medical visits, and blood test results from medical history. The dataset counts 9,471 medical examinations and 12,735 blood test results. For almost all of the parameters relevant for research on cardiovascular risk, data completeness and availability is over 90 %, but it varies with time and improves in the latest time period. In the absence of biobanks, collecting and computerising available good-quality occupational medicine archive data constitutes a valuable alternative for epidemiological and aetiological research in occupational health. Biobanks rarely contain biological samples over an entire worker's career and medical data from nuclear industry archives might make up for unavailable biomarkers that could provide information on cardiovascular and metabolic diseases.

KEY WORDS: *aetiology; cardiovascular mortality; epidemiology; ionising radiation; occupational medicine*

Studies able to properly assess the risk of internal contamination with radioactive compounds are scarce. This can partly be explained by the difficulty to reliably reconstruct the received internal dose or by the need of a long and robust follow-up to be able to do that. In France, the population of nuclear workers from AREVA NC's Pierrelatte facility appears to be interesting in this respect. The main activities at this plant involved uranium hexafluoride (UF₆) enrichment between 1959 and 1990 and the chemical conversion of uranium in the 1980s. The plant produced uranium compounds enriched in ²³⁵U to different degrees from

natural uranium ore but also reprocessed uranium from spent fuel turned into stable compounds for storage. Some workers were therefore potentially exposed to inhalation of uranium particles and other industrial chemicals. A retrospective pilot study in 2005 aimed at examining the effects of long-term, low-dose exposure to different uranium compounds in a cohort of plant workers (1). In the 2010 extension (2), the cohort totalled 2,897 workers who remained at the plant for at least six months between 1 January 1968 and 31 December 2006. At first, the available information regarding these workers could not be used

to reconstruct individual occupational exposure to uranium. For this reason, an epidemiological approach consisting of a semi-quantitative exposure assessment was developed in 2007 (3) and later a specific job exposure matrix (JEM) was designed to obtain a retrospective estimate of cumulative exposure to different uranium compounds and other pollutants (4). On the basis of this exposure assessment a relationship was established for lung and hematopoietic cancer (5).

Over the last decade, a potential risk of cardiovascular diseases associated with ionising radiation has become a major issue (6-11) but with inconsistent results in low-dose exposure studies (12, 13). Therefore, possible cardiovascular effects of internal uranium contamination were also studied in the Pierrelatte workers cohort (2). The results suggest that exposure to slowly soluble uranium compounds, notably reprocessed uranium compounds, may increase the risk of mortality from circulatory system diseases. The main limitations to analysing this cohort with specific internal exposure was the absence of individual biological and lifestyle data, especially about the known cardiovascular risk factors as defined by the Framingham study (14-16) or the international INTERHEART study (17, 18). Another limitation was the absence of organ-specific doses from uranium exposure.

Occupational medicine (OM) in French nuclear industry began simultaneously with uranium processing at the beginning of the 1950s. The principal aim of occupational physicians was to detect workers' health problems at the time of employment and to monitor their health status over the entire professional life. In this new domain, the first guidelines were mainly based on the experience of physicians and medical research pioneers. Thanks to national regulations requiring that occupational medical records are archived, most cardiovascular risk parameters are still kept on record, covering a period from 1958 to the present.

We studied these OM archives and collected the basic anthropometric data (height, weight), blood pressure (systolic and diastolic), and blood tests for all the monitored workers. Our intention was to test the hypothesis that low-dose ionising radiation increases the risk of cardiovascular diseases, taking into account these individual parameters. For this reason we have launched a nested case-control study that includes 394 workers (111 cases, up to 5 controls per case) from the cohort. The study is still under way.

In this preliminary report, however, we focus on the dataset collected in the above study in order to see whether the collected data could be useful for an epidemiological study or an aetiological research and to discuss other potential uses and limitations of OM records such as this. We analysed the type, quality, and variations in clinical and recording practice over time.

MEDICAL DATA COLLECTION AND ANALYSIS

Study population

The data presented here are related to a subsample of 394 eligible workers from the cohort of 2,897 workers employed at the AREVA NC Pierrelatte plant for at least six months between 1960 and 31 December 2006. The larger cohort was analysed for cancer and cardiovascular outcomes with unexpected results (2, 5) that called for further investigation. We collected additional data on potential risk factors for cardiovascular and metabolic diseases in order to establish their relationship with long-term exposure to low-dose uranium. In the choice of data to collect we primarily relied on the INTERHEART study (17, 18). The subsample was obtained by matching 111 cases of death from cardiovascular diseases (CVD) with controls according to sex, calendar period, five-year age classes, and socio-economic status at hire (clerical workers, engineering / management or technician) to account for potential confounding, especially regarding uranium exposure data which were available from the cohort study (19). Each case was matched to up to five controls whenever possible. One control could match more than one case, if the matching criteria were met. People who later died but at a certain earlier point met these matching criteria were also used as controls. To obtain the sample of controls who were alive at the time of death of the case we used an incidence density sampling method (20). Seventy cases were matched to five controls, five to four controls, another five to three controls, nine to two controls, and 13 cases to one control. No match was found for nine cases, including two women cases and cases with unmatchable combination of age and socio-economic status at hire.

Monitoring time frame and database creation

The data analysed here have been collected from OM monitoring records that span from 1959 to 2011, starting from workers' screening at hire to the last follow-up. Medical examination by occupational physicians was annual, but some of the workers received bi- or trimestral check-up, depending on exposure conditions.

Most of the paper archives were found in the Occupational Medicine Department (OMD) of the AREVA NC Pierrelatte plant. When the last employment was in another plant of the AREVA Group, we were able to track down the medical file of the worker. The files matching our population set were coded by ID numbers.

For each worker we reconstructed a computer file containing full medical records by manually entering all data from the paper archive. All results had been converted to the more recent unit. The data were then split into 14 interrelated database tables: *identification* (including the first and last day of work at the plant, height, and full medical history); *previous employments*; *job positions*; *job description*; *smoking* (describing changes in tobacco consumption); *pulmonary radiographs*; *external dosimetry*; *radio toxicological analyses results*; *accidental contaminations*; *number and causes of sick leave*; *capability certificates*; *electrocardiogram reports*, *medical examination findings*, and *blood tests*. The last two tables were the subject of our analysis, as the parameters contained in them provide key information about CVD risks. The medical examination table includes data about weight, systolic and diastolic blood pressure, pulse, and cardiovascular, cancer, or other disease observations registered at each visit. The blood tests include standard haematological parameters (erythrocytes, leucocytes, haematocrit, basophils, eosinophils, lymphocytes, monocytes, and platelets) and available metabolic parameters (total cholesterol, high-density lipoprotein (HDL), low-density lipoprotein (LDL), urea, blood sugar, triglycerides, creatinine, urea and uric acid, and gamma-glutamyl-transpeptidases (GGT)).

For all workers we computed body mass index (BMI) based on the standard formula:

$$BMI (kg / m^2) = \frac{weight}{height^2}$$

For the body surface area (BSA) we used the Dubois and Dubois formula (21):

$$BSA = 0.20247 * Height(m)^{0.725} * weight(kg)^{0.425}$$

Data quality and analysis

The quality of data in the database tables was assessed primarily by checking maximum and minimum values to identify and correct extreme values. Then the data of 30 randomly selected paper files were systematically reviewed to identify potential mistakes. All database values were compared to paper files. Of nearly 1,700 database values, only nineteen were wrong, indicating an error rate of about 1 %. These errors were further checked by comparing them with individual parameter means and medians and normal ranges.

For all parameters of interest we calculated two main indicators: availability and completeness. Availability is the percentage of workers with at least one recorded value and completeness is the number of recorded medical examination or blood test results divided by the total number of workers.

RESULTS AND DISCUSSION

The main characteristics of the collected data are given in Table 1. Among the workers, 12 are still

Table 1 Main characteristics of collected data at the AREVA NC Pierrelatte uranium processing plant

Characteristics	N
Population	394 (100 %)
Potentially exposed to uranium compounds	330 (84 %)
Female workers	9 (2 %)
Mean age at recruitment (y) (min-max)	35 (20-57)
Median age at recruitment (y)	35.3
Mean OM* follow-up (y) (min-max)	21 (1-37)
Median OM* follow-up (y)	20.2
Total number of medical examinations	9471
Mean number of medical examinations per worker (min-max)	25 (1-57)
Median number of medical examination	22
Total number of results of blood test from medical history	12735
Mean number of results of blood test from medical history per worker (min-max)	32 (1-93)
Median number of results of blood test from medical history per worker	31

*OM=Occupational Medicine

employed. The mean time between the last medical examination and the occurrence of death by cardiovascular disease is 11.61 years (± 8 years). Workers with more than 10 years since the last medical examination make 56.8 % of the population.

Over the monitoring period, 9,471 medical examinations were performed and 12,735 blood test results obtained (Table 2). In terms of the means this translates to 25 medical examinations and 35 blood test results per person. The data show that 187 (47.10%) had 20-29 medical examinations and 286 (72.6%) at least 20 medical examinations. Only 34

(8.2%) workers have less than 20 blood test results recorded.

This database clearly shows that changes in the cardiovascular system (e.g. systolic and diastolic pressure) of the workers had been carefully monitored. The number of medical examinations and blood tests is sufficient to provide reliable information about a relevant parameter for further epidemiological or aetiological research.

Table 3 shows the completeness and the availability of specific variables of interest. It ranged between 1.2 % for HDL cholesterol and 99.8 % for granulocytes,

Table 2 Anthropometric and biomedical data collected at the AREVA NC Pierrelatte uranium processing plant for 394 workers

Variables	Unit	N of measurements	Range
<i>Medical examination (N=9471)</i>			
Height	cm	394 ⁱ	152-190
Weight	kg	9263	44-116
Systolic blood pressure	mmHg	9379	90-290
Diastolic blood pressure	mmHg	9375	40-170
Body Mass Index	kg m ⁻²	9256	16.6-37.4
Body Surface Area	m ²	9206	1.4-2.4
<i>Blood tests from medical history (N=12735)</i>			
Erythrocytes count	million mm ⁻³	12696	4.1-6.0
Haemoglobin	g dL ⁻¹	12480	3.64-17.8
Haematocrit	%	11639	33.3-62.0
Mean corpuscular volume	m ³	11932	32.2-199.0
Leucocytes	n mm ⁻³	12696	4000-25600
Lymphocytes	%	12709	0-78
Monocytes	%	12709	0-56
Thrombocytes	n mm ⁻³	2445	91-685
Neutrophil granulocytes	%	12709	19-88
Eosinophil granulocytes	%	12709	0-36
Basophil granulocytes	%	12709	0-7
<i>Metabolic parameters</i>			
Total cholesterol	g L ⁻¹	7284	1.0-5.1
LDL ⁱⁱ cholesterol	g L ⁻¹	993	0.2-2.4
HDL ⁱⁱⁱ cholesterol	g L ⁻¹	153	0.4-3.6
Glycaemia	g L ⁻¹	7169	0.1-3.3
Triglycerides	g L ⁻¹	1273	0.3-12.5
Creatinine	mg L ⁻¹	1623	0.2-19.0
GGT	IU L ⁻¹	1452	2-112.5
Urea	g L ⁻¹	11385	0.1-0.47
Uric acid	mg L ⁻¹	7169	0-140

ⁱHeight was collected once for each worker

ⁱⁱLDL= low density lipoproteins

ⁱⁱⁱHDL=high density lipoproteins

Table 3 *Completeness and availability of worker follow-up data from medical examinations and blood tests*

Variables	Completeness of data*	Availability of workers' data**	Year of first and last recorded value
<i>Medical visits</i>			
Height	99.8 %	99.8 %	1959-2011
Weight	97.7 %	99.2 %	1959-2011
Diastolic blood pressure	99.0 %	99.2 %	1959-2011
Systolic blood pressure	98.9 %	99.2 %	1959-2011
Pulse	90.4 %	99.2 %	1959-2011
BMI/BSA	97.6 %	99.2 %	1959-2011
<i>Results of blood test from medical history</i>			
Erythrocytes	99.7 %	99.2 %	1959-2011
Haemoglobin	98.0 %	99.2 %	1959-2011
Haematocrits	91.4 %	98.7 %	1959-2011
Mean corpuscular volume	93.7 %	99.2 %	1959-2011
Leucocytes	99.7 %	99.2 %	1959-2011
Lymphocytes	99.8 %	99.2 %	1959-2011
Monocytes	99.8 %	99.2 %	1959-2011
Thrombocytes	19.2 %	92.0 %	1959-2011
Neutrophil granulocytes	99.8 %	99.2 %	1959-2011
Eosinophil granulocytes	99.8 %	99.2 %	1959-2011
Basophil granulocytes	99.8 %	99.2 %	1959-2011
<i>Metabolic parameters</i>			
Total cholesterol	57.2 %	98.5 %	1961-2011
LDL cholesterol	7.8 %	30.0 %	1985-2011
HDL cholesterol	1.2 %	8.3 %	1968-2011
Glycaemia	56.3 %	98.5 %	1959-2011
Triglycerides	10.0 %	42.3 %	1969-2011
Urea	89.4 %	99.0 %	1959-2011
Creatinine	12.7 %	69.3 %	1966-2011
GGT	11.4 %	40.8 %	1963-2011
Uric acid	56.3 %	98.5 %	1965-2011

*Percentage calculated as the number of collected values over the number of medical examination or results of blood test from the medical history for each workers, divided by the total number of workers with collected values

**Percentage of workers with at least one value

lymphocytes, monocytes, and height. The availability of specific variables ranged from 8.3 % for HDL cholesterol to 99.8 % for height. The high completeness of blood count data provides reliable information for investigating possible effects of exposure to ionising radiation and uranium on blood count.

Change of recording practice over time

Workers hired before 1970 make 91 % of our population. For them, the completeness of almost every parameter of medical examination or blood test is $\geq 90\%$. The exceptions are HDL and LDL cholesterol, triglyceride, GGT, and creatinine. Until the 1960s only

high cholesterol values were recorded. Routine recording started only after 1973, with the first major publications on total cholesterol as a cardiovascular risk factor (22-25). For workers hired later than 1970 the completeness rose highly, except, once again for HDL cholesterol (Figure 1).

The choice of parameters and their biological relevance

The sample analysed in this study provided highly complete and available data on the risk factors for acute myocardial infarction, including the history of hypertension (almost 99 % completeness for systolic

and diastolic pressure), abdominal obesity (by calculating BMI/BSA), total cholesterol, and diabetes (using blood sugar as an indicator). About 50 % of the workers had both blood sugar and cholesterol measured. Records on psychosocial risk factors like stress and hard physical activity at work or physical activity in free time were too scarce or imprecise to use as required by the INTERHEART study (17).

Since the follow-up was no more compulsory after retirement, few data cover that period. However, because most workers were exposed to chronic low-rate exposure to ionising radiation and uranium rather than acute, high-rate accidental exposure, potential health effects would be mainly long-term, evolving from light to moderate biological and physiological changes that could lead to clinically detectable symptoms or a disease once they become irreversible. Therefore, it should be possible to identify any changes in parameters from the data collected during employment and investigate how these changes may influence the occurrence of cardiovascular disease from a mechanistic point of view.

One major advantage of our approach is that for most of the parameters it can provide not only one value per individual, but the whole variation over time. For example, every worker had a mean of 23 systolic blood pressure measurements over the median follow-up of 20 years.

The Framingham Cardiac Risk Score (22) is a well-known tool to predict 10-year cardiovascular risks for an individual. It requires seven variables to calculate the risk: sex, age, total cholesterol, HDL cholesterol, systolic and diastolic blood pressure, diabetic status, and smoking status. The availability of data for all the required parameters in our dataset

is good, with the notable exception of HDL cholesterol (which is available for only 33 workers). The diabetic status can be determined on the basis of blood sugar levels. Smoking records are not uniform (detailed information on tobacco consumption and duration of smoking is only partly known) and can only be analysed as a binary variable across the sample.

Considering that blood count parameters in our sample have almost 100 % completeness and availability over time, it would be relatively easy to evaluate the effects of low-dose ionising radiation on blood counts (23, 24). Diastolic and systolic blood pressure can also be analysed in relation to internal uranium dose received before retirement. With the added benefit of blood counts and medical examination parameters, robust studies are possible.

In contrast, the majority of metabolic parameters reached 50 % completeness in the 1980's and could therefore be considered only with great caution and an extension of the workers' follow-up.

The only important variables with low availability in our dataset are creatinine and HDL cholesterol. Creatinine is a marker of renal function, fundamental in uranium internal contamination, while HDL is an important element of the Framingham score.

Potential use in the upcoming nested case-control study

The collected data will be used in the upcoming nested case-control study as co-variables in addition to the main exposure variables (e.g. exposure to ionising radiation and uranium compounds) to investigate how they could have affected the results of the previous cohort study by Guseva Canu et al.

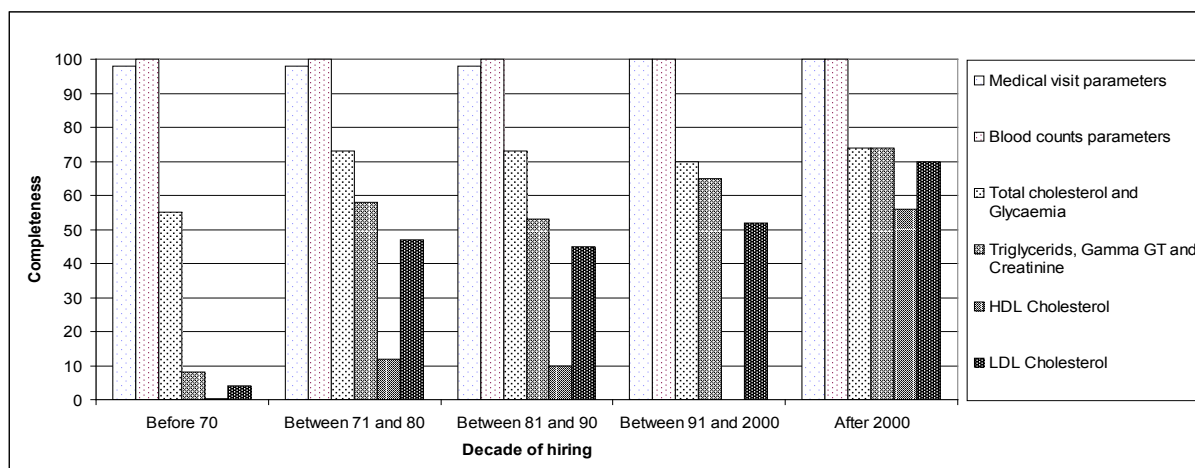


Figure 1 Completeness of medical examination and blood test parameters data by decades

(2). In that study, the effects of exposure to uranium compounds on cardiovascular mortality were assessed without taking into account the known cardiovascular risk factors, as they were not available at the time. Therefore, the collected variables will primarily be analysed as possible confounding or modifying factors using a conditional logistic model. These analyses should be able to handle time-dependent covariates because of the longitudinal changes in blood tests and anthropometry over time.

Potential for broader use

The large set of collected haematological data will allow us to address the hypotheses on the association between haematological parameters/hypertension and internal uranium contamination. For instance, Wagner et al. (25) found that exposed women ran a higher risk of elevated systolic pressure.

The use of longitudinal data makes it possible to evaluate correlations within and between workers using generalised estimating equations (26). Thus, in addition to assessing the modifying or confounding effects on cardiovascular mortality, OM data can also be used independently, as an intermediate health outcome to assess the effect of uranium exposure on health parameters such as blood pressure.

In that sense, OM data could also be used for aetiological research. Kathren et al. (27, 28) have already pointed to acute internal exposure to uranium compounds as the possible cause of kidney disease (27, 28). Low-dose radiation effects are less consistent in the literature, but some of the evidence calls for further studies (29, 30). In fact, the archived files we reviewed do refer to adverse renal effects, and consistently so to renal colic. We believe that data on creatinine will be useful for pending studies of the effects of radiation on the renal function in the plant workers.

Furthermore, the OM data appear to be an important source of information for retrospective studies. A longitudinal follow-up of workers makes possible the use of various epidemiological and statistical models to retrospectively assess exposure risk and build a robust prospective cohort. If no specific sampling of biological material is available, the use of routinely collected OM data may be justified prior to further research. A major limitation to the use of OM data for epidemiological research are the missing or incorrect values in the database. However, this issue can be addressed with frequentist (31) or Bayesian (32) or other methods (33). If there is no

pattern to missing data, this could be resolved by careful data management, but if the errors are systematic, Bayesian method could help to fill them in.

As biobanks are very difficult to set up in France for ethical and economic reasons, the electronic storage of half a century of archived OM data on nuclear workers is a cost-effective alternative. Such data are generally directly available, accurate, and free of recall bias. Most are systematic (annual) biological measurements that span the worker's career and follow strict medical surveillance procedures. Moreover, biobanks rarely keep biological samples for an entire career.

Finally, since the OM follow-up provides epidemiologists with longitudinal data, it also provides an insight into the causal relationship between exposure to ionising radiation and symptomatic coronary artery disease through the effects of radiation on leucocyte and monocyte counts (34). The use of data already obtained by health monitoring should efficiently complement molecular epidemiology studies. These data could be used as a reference for additional data collected after retirement. Furthermore, systematic blood records for each worker make it easier to collect additional samples to analyse new biomarkers (DNA, mutations, epigenetic markers, etc...).

Overall, the recording practices at the AREVA NC Pierrelatte facility provide a reliable basis for epidemiological studies, especially regarding blood counts and medical examination parameters. The completeness and availability of the data are relatively high for known major cardiovascular risk factors. The major strength of our subsample database is that it can reduce/control potential confounding and effect-modifying factors. Data on the full cohort will allow an even more precise evaluation of radiation effects by taking into account the parameters that can affect the circulatory system.

Acknowledgements

The authors wish to thank Olivier Laurent from the Laboratoire d'épidémiologie for carefully reading the drafts, Iris Jovanovic for her major role in creating the computer database, the personnel of the Pierrelatte Occupational Medicine Department for great help in logistics, and Dr Alain Acker (AREVA) for his help in accessing the workers' medical records in the Pierrelatte OM department.

Conflict of interest

The authors declare no conflict of interest.

REFERENCES

- Guseva Canu I, Cardis E, Metz-Flamant C, Caër-Lorho S, Auriol B, Wild P, Laurier D, Tirmarche M. French cohort of the uranium processing workers: mortality pattern after 30-year follow-up. *Int Arch Occup Environ Health* 2010;83:301-8. doi: 10.1007/s00420-009-0455-0
- Guseva Canu I, Garsi JP, Caër-Lorho S, Jacob S, Collomb P, Acker A, Laurier D. Does uranium induce circulatory diseases? First results from a French cohort of uranium workers. *Occup Environ Med* 2012;69:404-9. doi: 10.1136/oemed-2011-100495
- Guseva Canu I, Molina G, Goldberg M, Collomb P, David JC, Perez P, Paquet F, Tirmarche M. Construction d'une matrice emplois-expositions pour le suivi épidémiologique des travailleurs de l'industrie nucléaire en France [Development of a job exposure matrix for the epidemiological follow-up of workers in the French nuclear industry, in French]. *Rev Epidemiol Sante Publique* 2008;56:21-9. doi: 10.1016/j.respe.2007.11.001
- Guseva Canu I, Paquet F, Goldberg M, Auriol B, Bérard P, Collomb P, David JC, Molina G, Perez P, Tirmarche M. Comparative assessing for radiological, chemical, and physical exposures at the French uranium conversion plant: Is uranium the only stressor? *Int J Hyg Environ Health* 2009;212:398-413. doi: 10.1016/j.ijheh.2008.09.002
- Guseva Canu I, Jacob S, Cardis E, Wild P, Caër S, Auriol B, Garsi JP, Tirmarche M, Laurier D. Uranium carcinogenicity in humans might depend on the physical and chemical nature of uranium and its isotopic composition: results from pilot epidemiological study of French nuclear workers. *Cancer Causes Control* 2011;22:1563-73. doi: 10.1007/s10552-011-9833-5
- Azizova TV, Muirhead CR, Druzhinina MB, Grigoryeva ES, Vlasenko EV, Sumina MV, O'Hagan JA, Zhang W, Haylock RG, Hunter N. Cerebrovascular diseases in the cohort of workers first employed at Mayak PA in 1948-1958. *Radiat Res* 2010;174:851-64. doi: 10.1667/RR1928.1
- Azizova TV, Muirhead CR, Druzhinina MB, Grigoryeva ES, Vlasenko EV, Sumina MV, O'Hagan JA, Zhang W, Haylock RG, Hunter N. Cardiovascular diseases in the cohort of workers first employed at Mayak PA in 1948-1958. *Radiat Res* 2010;174:155-68. doi: 10.1667/RR1789.1
- Little MP, Azizova TV, Bazyka D, Bouffler SD, Cardis E, Chekin S, Chumak VV, Cucinotta FA, de Vathaire F, Hall P, Harrison JD, Hildebrandt G, Ivanov V, Kashcheev VV, Klymenko SV, Kreuzer M, Laurent O, Ozasa K, Schneider T, Tapio S, Taylor AM, Tzoulaki I, Vandoolaeghe WL, Wakeford R, Zablotska LB, Zhang W, Lipshultz SE. Systematic review and meta-analysis of circulatory disease from exposure to low-level ionizing radiation and estimates of potential population mortality risks. *Environ Health Perspect* 2012;120:1503-11. doi: 10.1289/ehp.1204982
- McGale P, Darby SC. Low doses of ionizing radiation and circulatory diseases: a systematic review of the published epidemiological evidence. *Radiat Res* 2005;163:247-57. PMID: 15733031
- Little MP, Gola A, Tzoulaki I. A model of cardiovascular disease giving a plausible mechanism for the effect of fractionated low-dose ionizing radiation exposure. *PLoS Comput Biol* 2009;5:e1000539. doi:10.1371/journal.pcbi.1000539
- McGeoghegan D, Binks K, Gillies M, Jones S, Whaley S. The non-cancer mortality experience of male workers at British Nuclear Fuels plc, 1946-2005. *Int J Epidemiol* 2008;37:506-18. doi: 10.1093/ije/dyn018
- Metz-Flamant C, Bonaventure A, Milliat F, Tirmarche M, Laurier D, Bernier MO. Irradiations à faibles doses et risque de pathologie cardiovasculaire: revue des études épidémiologiques [Low doses of ionizing radiation and risk of cardiovascular disease: A review of epidemiological studies, in French]. *Rev Epidemiol Sante Publique* 2009;57:347-59. doi: 10.1016/j.respe.2009.04.009
- Little MP. Cancer and non-cancer effects in Japanese atomic bomb survivors. *J Radiol Prot* 2009;29(2A):A43-59. doi: 10.1088/0952-4746/29/2A/S04
- Kannel WB, Castelli WP, Gordon T, McNamara PM. Serum cholesterol, lipoproteins, and the risk of coronary heart disease. The Framingham study. *Ann Intern Med* 1971;74:1-12. PMID: 5539274
- Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J. Factors of risk in the development of coronary heart disease - six year follow-up experience. The Framingham Study. *Ann Intern Med* 1961;55:33-50. PMID: 13751193
- Kannel WB, McGee D, Gordon T. A general cardiovascular risk profile: the Framingham Study. *Am J Cardiol* 1976;38:46-51. PMID: 132862
- Rosengren A, Hawken S, Ounpuu S, Sliwa K, Zubaid M, Almahmeed WA, Blackett KN, Sitthi-amorn C, Sato H, Yusuf S; INTERHEART investigators. Association of psychosocial risk factors with risk of acute myocardial infarction in 11119 cases and 13648 controls from 52 countries (the INTERHEART study): case-control study. *Lancet* 2004;364:953-62. doi: 10.1016/S0140-6736(04)17019-0
- Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, McQueen M, Budaj A, Pais P, Varigos J, Lisheng L; INTERHEART Study Investigators. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. *Lancet* 2004;364:937-52. doi: 10.3410/f.1097404.553463
- Breslow NE, Day NE, editors. Statistical methods in cancer research. Volume I - The analysis of case-control studies. IARC Sci Publ 1980;32:5-338.
- Richardson DB. An incidence density sampling program for nested case-control analyses. *Occup Environ Med* 2004;61(12):e59.
- DuBois D, DuBois EF. A formula to estimate the approximate surface area if height and weight be known. *Arch Intern Med* 1916;17:863-71.
- Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97:1837-47. doi: 10.1161/01.CIR.97.18.1837
- Shawky S, Amer HA, Hussein MI, el-Mahdy Z, Mustafa M. Uranium bioassay and radioactive dust measurements at

- some uranium processing sites in Egypt - health effects. *J Environ Monit* 2002;4:588-91. PMID:12196006
24. Akushevich IV, Veremeyeva GA, Dimov GP, Ukraintseva SV, Arbeev KG, Akleyev AV, Yashin AI. Modeling deterministic effects in hematopoietic system caused by chronic exposure to ionizing radiation in large human cohorts. *Health Phys* 2010;99:322-9. doi: 10.1097/HP.0b013e3181c61dc1
25. Wagner SE, Burch JB, Bottai M, Pinney SM, Puett R, Porter D, Vena JE, Hébert JR. Hypertension and hematologic parameters in a community near a uranium processing facility. *Environ Res* 2010;110:786-97. doi: 10.1016/j.envres.2010.09.004
26. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics* 1988;44:1049-60. PMID: 3233245
27. Kathren RL, Burklin RK. Acute chemical toxicity of uranium. *Health Phys* 2008;94:170-9. doi: 10.1097/01.HP.0000288043.94908.1f
28. Kathren RL, McInroy JF, Moore RH, Dietert SE. Uranium in the tissues of an occupationally exposed individual. *Health Phys* 1989;57:17-21. PMID: 2745077
29. Thun MJ, Baker DB, Steenland K, Smith AB, Halperin W, Berl T. Renal toxicity in uranium mill workers. *Scand J Work Environ Health* 1985;11:83-90. PMID: 3890163
30. Arzuaga X, Rieth SH, Bathija A, Cooper GS. Renal effects of exposure to natural and depleted uranium: a review of the epidemiologic and experimental data. *J Toxicol Environ Health B Crit Rev* 2010;13:527-45. PMID: 21170808
31. Soullier N, de La Rochebrochard E, Bouyer J. Multiple imputation for estimation of an occurrence rate in cohorts with attrition and discrete follow-up time points: a simulation study. *BMC Med Res Methodol* 2010;10:79. doi: 10.1186/1471-2288-10-79
32. Ibrahim JG, Chen MH, Kim S. Bayesian variable selection for the Cox regression model with missing covariates. *Lifetime Data Anal* 2008;14:496-520. doi: 10.1007/s10985-008-9101-5
33. Ibrahim JG, Zhu H, Tang N. Model Selection Criteria for Missing-Data Problems Using the EM Algorithm. *J Am Stat Assoc* 2008;103:1648-58. PMID: 19693282
34. Advisory Group on Ionizing Radiation (AGIR). Potential mechanism of cardiovascular injury and their relevance to radiation, in circulatory disease risk: report of the independent. London: HPA; 2010.

Sažetak**Polustoljetni arhiv odjela za medicinu rada francuske nuklearne elektrane - prašnjavo skladište ili zlatni rudnik za epidemiološka istraživanja?**

U ovome se članku raspravlja o dostupnosti i potpunosti medicinskih podataka o radnicima iz nuklearne elektrane AREVA NC Pierrelatte te o njihovoj korisnosti za epidemiološka ispitivanja krvožilnih i metaboličkih poremećaja povezanih s izloženosti uranu. Na temelju zdravstvenih kartona 394 radnika koji su dio većeg kohortnog ispitivanja s 2897 radnika u francuskim nuklearnim elektranama stvorena je elektronička baza podataka s informacijama o ranijem zaposlenju, položaju, opisu posla, liječničkim pregledima i krvnim nalazima. Baza obuhvaća podatke iz 9.471 liječničkog pregleda te 12.735 krvnih nalaza. Dostupnost i potpunost podataka za gotovo sve parametre važne za procjenu rizika od bolesti srca i krvožilja bile su više od 90 %, ali neujednačene, i popravljale su se tek prema kraju 50-godišnjeg praćenja. U nedostatku biobanaka, prikupljanje i digitalizacija pouzdanih arhivskih podataka iz medicine rada vrijedan su alternativni izvor za epidemiološka i etiološka ispitivanja. Biobanke rijetko čuvaju biološke uzorke radnika cijeli njegov radni vijek, pa zdravstveni podaci čuvani u arhivima nuklearnih elektrana mogu kvalitetno dopuniti prazninu uslijed nedostatka biomarkera i dati uvid u nastanak i razvoj bolesti metabolizma te srca i krvožilja.

KLJUČNE RIJEČI: *bolesti srca i krvožilja; epidemiologija; etiologija; ionizirajuće zračenje; medicina rada; smrtnost*

CORRESPONDING AUTHOR:

Dominique Laurier
IRSN Institut de Radioprotection et de Sûreté Nucléaire
Laboratoire d'Epidémiologie des Rayonnements ionisants
PRP-HOM/SRBE/LEPID
BP 17 92262 Fontenay aux Roses Cedex, France
E-mail: dominique.laurier@irsn.fr