Different approaches of selecting the underlying causes of death across Russian regions

Inna Danilova [1,2], Dmitry Jdanov [1,3], Vladimir Shkolnikov [1,3]

[1] Max Planck Institute for Demographic Research, Rostock (Germany);
[2] National Research University Higher School of Economics, Moscow (Russia);
[3] The New Economic School, Moscow (Russia)

INTRODUCTION

Russian system for production of information on causes of death (CoDs) is decentralized. That may lead to inconsistency of coding practices within the country, biases in analyses of mortality by CoDs across regions, and lowering quality of CoD data at all-country level.

The present study assesses comparability of cause-specific mortality data across Russian regions. We examined the regional cause-specific data series as they were officially published and tried to measure regional discordances in propensity of coding various diagnoses. That allowed us making an indirect estimation of the uniformity of cause-of-death coding practices across the Russian territory and identify the most problematic causes of death and regions.

DATA

Death counts and population exposures were obtained from the Russian Federal State Statistics Service

Time period: 2002-2012

Territorial coverage:

- regions with the average annual population ≥ 1 million;
- excluding Chechen Rep. (death counts are unavailable before 2006);

→ A set of 52 regions (88.4% of the whole population, 88.5% of all deaths)

Causes of Death examined:

70 major groups of causes of death from the Russian Abridged Classification based on ICD-10

METHODS

As an indicator of cause-specific mortality prevalence, we used the cause-specific share of the all-cause age-standardized death rate:

$$S_{r,c,t} = \frac{SDR_{r,c,t}}{SDR_{r,t}} \cdot 100\% \quad ,$$

where $SDR_{r,c,t}$ is the age-standardized death rate for cause *c* in region *r* and year *t*, $SDR_{r,t}$ is the all-cause age-standardized death rate in region *r* and year *t*.

For each possible combination region/cause we had calculated the deviation from the cross-regional mean (period average):

$$V_{r,c} = \frac{1}{T}\sum_{t=1}^{T} \frac{|S_{r,c,t} - \overline{S_{\bullet,c,t}}|}{\overline{S_{\bullet,c,t}}} \cdot 100\% \quad ,$$

$$1 \le r \le 52, 1 \le c \le 70$$

(*A color visualization of the matrix* $\|V_{r,c}\|$ *presented in* ***Figure 1***)

where $\overline{S_{\bullet,c,t}}$ is the mean of regional $S_{r,c,t}$.

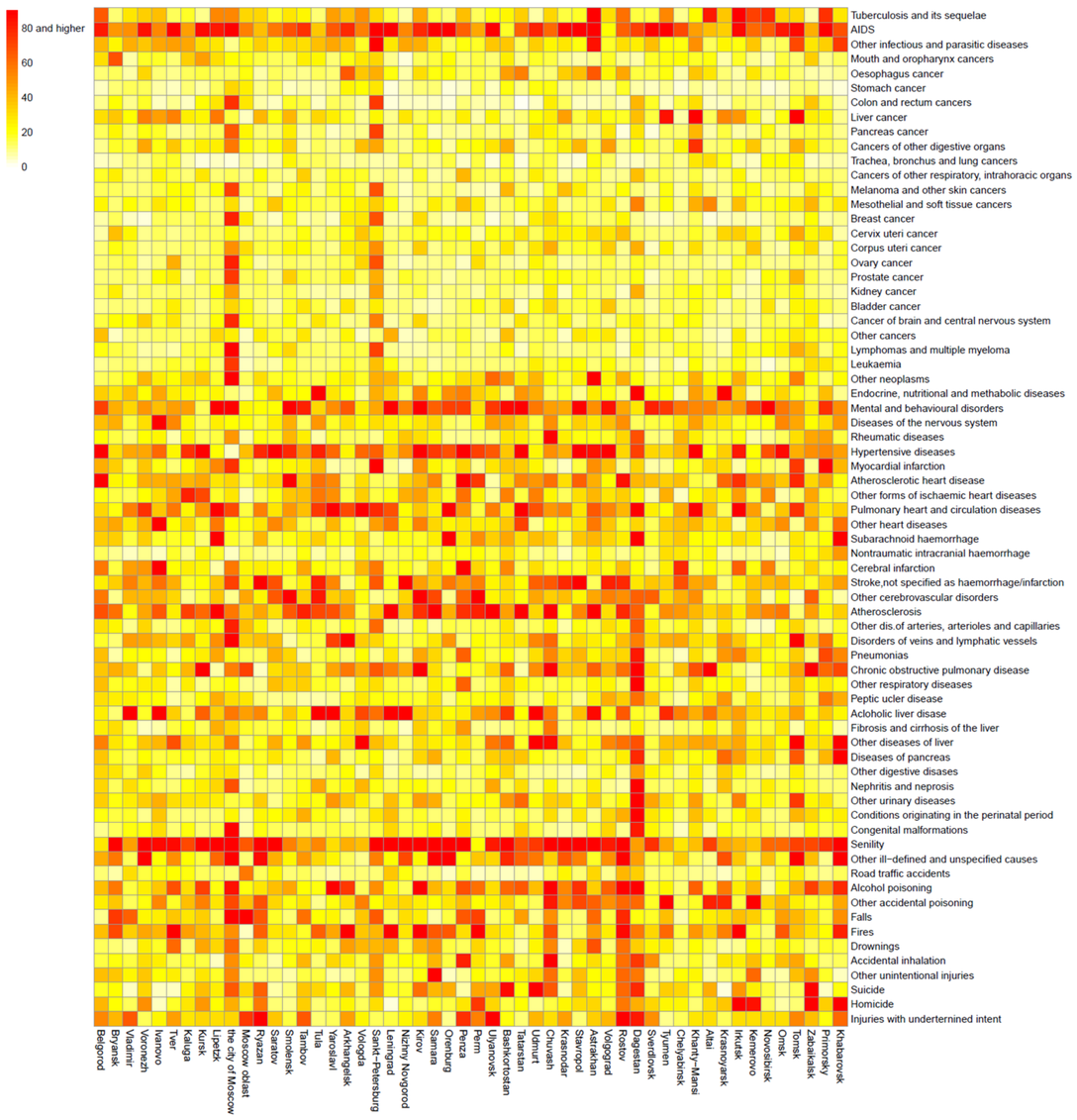
OLS regression for estimation of Region and CoD effects

$$V_{r,c} = a + b_r I_r + d_c I_c + \varepsilon_{r,c} \quad ,$$

where a is a constant term, $\varepsilon_{r,c}$ is an error term; I_r and I_c are the independent dummy variables for region and CoD; b_r and d_c are the coefficients on I_r and I_c.

RESULTS

Figure 1. Heatmap on inter-regional variability in causes of death. *The rows correspond to causes of death and the columns represent regions. The cells are colored according to* $V_{r,c}$ *values.*



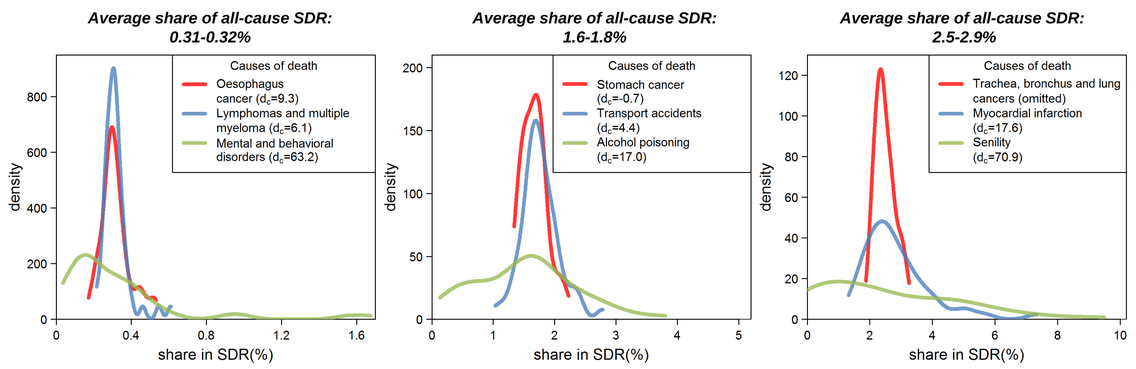
RESULTS

Table 1. Estimates of the regression coefficients of the OLS model (a fragment)

Cause of death	Coef. d_c (95% CI)	p-value	Region	Coef. b_r (95 CI)	p-value
Stomach cancer	**-0.67** (-10.94, 9.61)	0.90	Sverdlovsk	**-2.32** (-11.18, 6.54)	0.61
Trachea, bronchus and lung cancers	[Ref]	—	Chelyabinsk	**-2.31** (-11.17, 6.55)	0.61
Kidney cancer	**2.93** (-7.34, 13.21)	0.58	Novosibirsk	**-1.99** (-10.85, 6.87)	0.66
Breast cancer	**2.98** (-7.29, 13.26)	0.57	Moscow oblast	**-1.45** (-10.31, 7.41)	0.75
Colon and rectum cancer	**3.21** (-7.07, 13.49)	0.54	Omsk	**-0.87** (-9.73, 7.99)	0.85
⋮	⋮	⋮	⋮	⋮	⋮
Atherosclerotic heart disease	**34.02** (23.74, 44.3)	<0.001	Kaluga	[Ref]	—
Other symptoms and signs	**34.16** (23.88, 44.43)	<0.001	⋮	⋮	⋮
Other diseases of liver	**34.23** (23.95, 44.51)	<0.001	Astrakhan	**11.5** (2.64, 20.36)	0.01
Stroke not specified as hemorrhage or infarction	**36** (25.72, 46.28)	<0.001	Samara	**11.7** (2.84, 20.56)	0.01
Alcohol poisoning	**36.32** (26.04, 46.6)	<0.001	Rostov	**12.12** (3.26, 20.98)	0.01
Fires	**36.74** (26.46, 47.02)	<0.001	Tomsk	**12.53** (3.67, 21.39)	0.01
Alcoholic liver disease	**40.8** (30.52, 51.08)	<0.001	Lipetzk	**13.58** (4.72, 22.44)	<0.001
Chronic obstructive pulmonary dis.	**41.83** (31.55, 52.1)	<0.001	Chuvash	**16.74** (7.88, 25.6)	<0.001
Pulmonary heart disease and diseases of pulmonary circulation	**43.29** (33.01, 53.57)	<0.001	Sankt-Petersburg	**19.93** (11.07, 28.79)	<0.001
Hypertensive diseases	**51.4** (41.12, 61.68)	<0.001	city of Moscow	**29.78** (20.93, 38.64)	<0.001
Atherosclerosis	**53.76** (43.48, 64.04)	<0.001	Dagestan	**32.69** (23.83, 41.55)	<0.001
Mental and behavioral disorders	**63.14** (52.86, 73.41)	<0.001			
Senility	**70.88** (60.6, 81.16)	<0.001			
AIDS	**71.37** (61.09, 81.65)	<0.001			

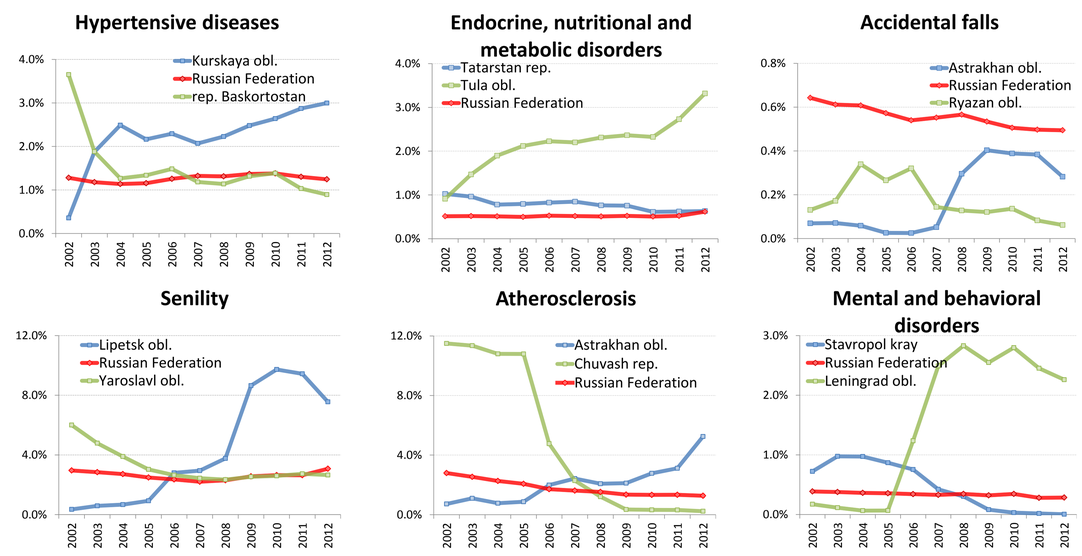
The estimated d_c coefficients reflect the horizontal patterns on the heatmap based on variation of cause-specific shares of the all-cause SDR across regions (*Figure 2*)

Figure 2. Selected distributions of cause-specific shares of the all-cause SDR across 52 regions. *Each panel presents three CoDs, which are close to each other by the share in the overall mortality with contrasting coefficients* d_c .



In addition to the cause-specific mortality structures across the regions of Russia, we also inspected the cause-specific trends across regions paying special attention to their volatility over time. A number of regional cause-specific series were unexpectedly distorted during the period 2002-2012 (***Figure 3***). These abrupt and/or large changes of mortality possibly indicate modifications of coding practices due to which some number of deaths that were previously coded to a certain item starts to be coded to the other one.

Figure 3. Examples of rapid and contrasting changes in regional cause-specific shares of all-cause SDR (both sexes combined). *The trend for Russia as a whole is provided for comparison.*



CONCLUSIONS

- The causes of death which can be easily diagnosed (malignant neoplasms, transport accidents, congenital malformations) experience roughly comparable cause-specific shares in regional mortality structures;
- The causes of death which can cause certain difficulties in defining whether they should be chosen as underlying (mental and behavioral disorders, diseases of nervous system, certain items inside the group of circulatory diseases) experience much higher inconsistency. Regional differences in mortality from these causes reflect peculiar coding practices rather than inform only about real mortality variation across regions;
- There is an intersection between the causes of death with a high spatial variability and the causes with a high volatility of regional trends;
- The prevalence of so-called “garbage” codes varies across Russian regions hugely. It affects regional mortality data for the other (“specified”) causes of death;
- For some causes of death, the concordance is higher when they are aggregated into broader diagnostic groups;
- Some regions introduce (or continue) very peculiar coding practices. This is probably related to a lack of centralized control.