



**Международная лаборатория
исследований населения и здоровья**

Семинар «Современная демография»

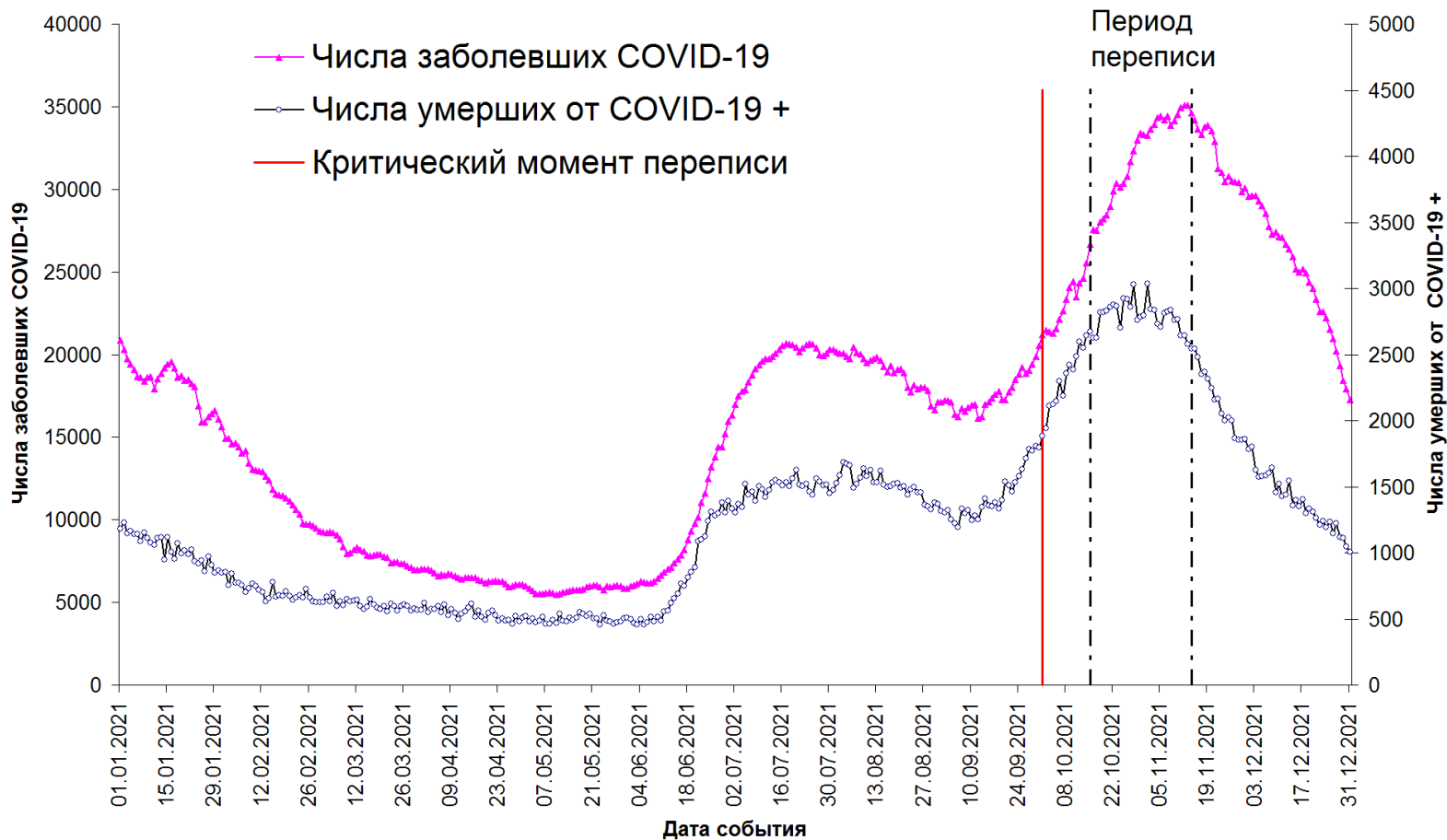
**О надежности и сопоставимости
результатов Всероссийской
переписи населения 2021 года**

Е.М. Андреев

8 июня 2023 г.

Всероссийская перепись населения России 2020 г. переписей уже тем, что прошла в период с 15 октября по 14 ноября 2021 г. и учитывала население по состоянию на 0 часов 01 октября 2021 г. Это означает, что все умершие с 1 октября по 14 ноября должны быть включены в перепись (по нашим расчетам их численность около 367 тыс. человек должны быть включены в переписные листы, а все родившиеся в этот период (примерно 174 тыс. человек) включены быть не должны.

Период переписи на фоне пандемии COVID-19 в России



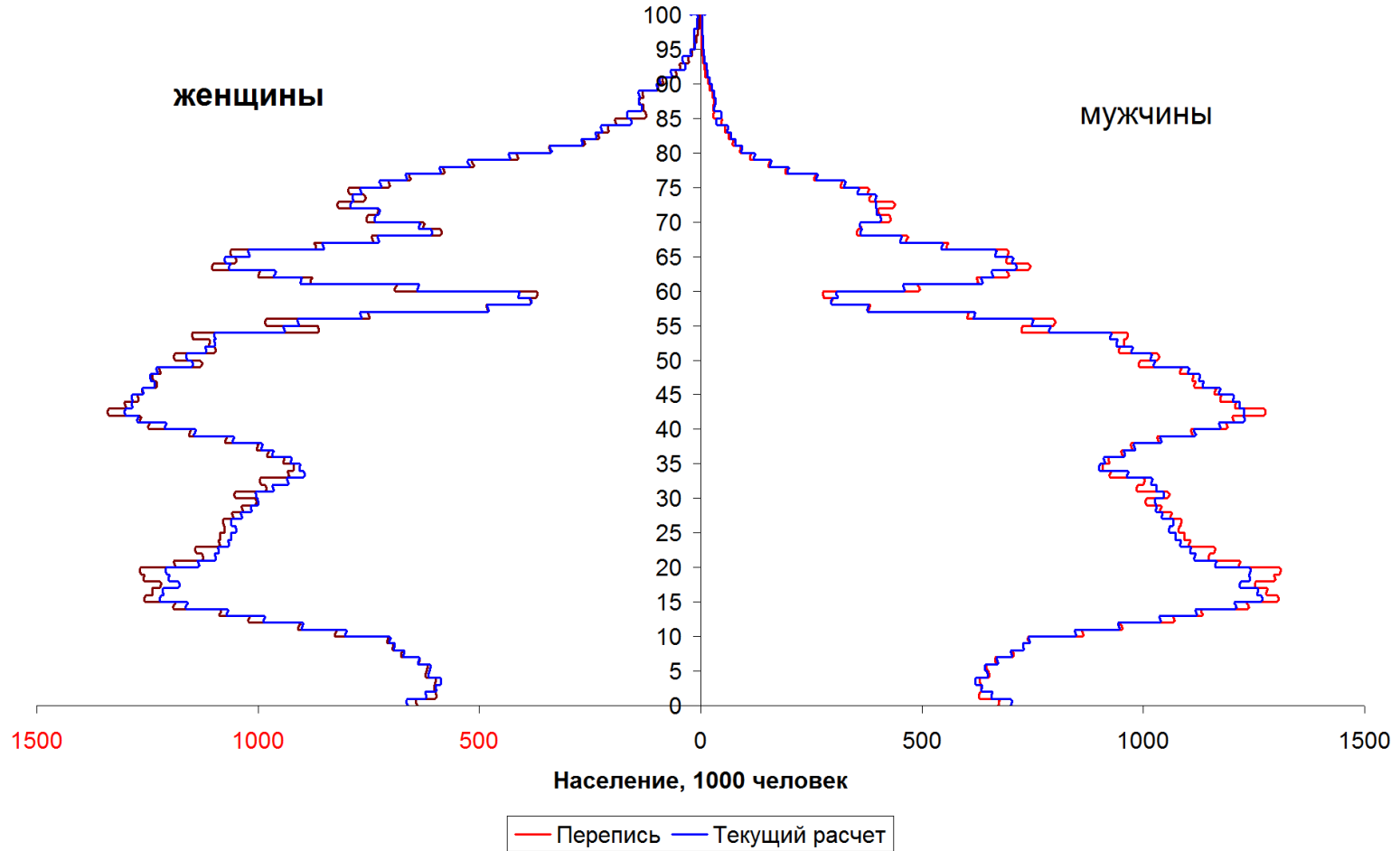
В какой мере особые условия проведения переписи повлияли на точность ее результатов?

Сравнение общей численности населения России по результатам переписей 2002*, 2010 и 2021 годов и согласно текущей оценке на момент переписи, тыс. человек

Год переписи	По данным переписи	Текущий расчет	Разность
А	1	2	1 - 2
2002	143943,8	142678,0	1266
2010	142817,1	141876,7	940
2021	147182,1	145712,2	1470

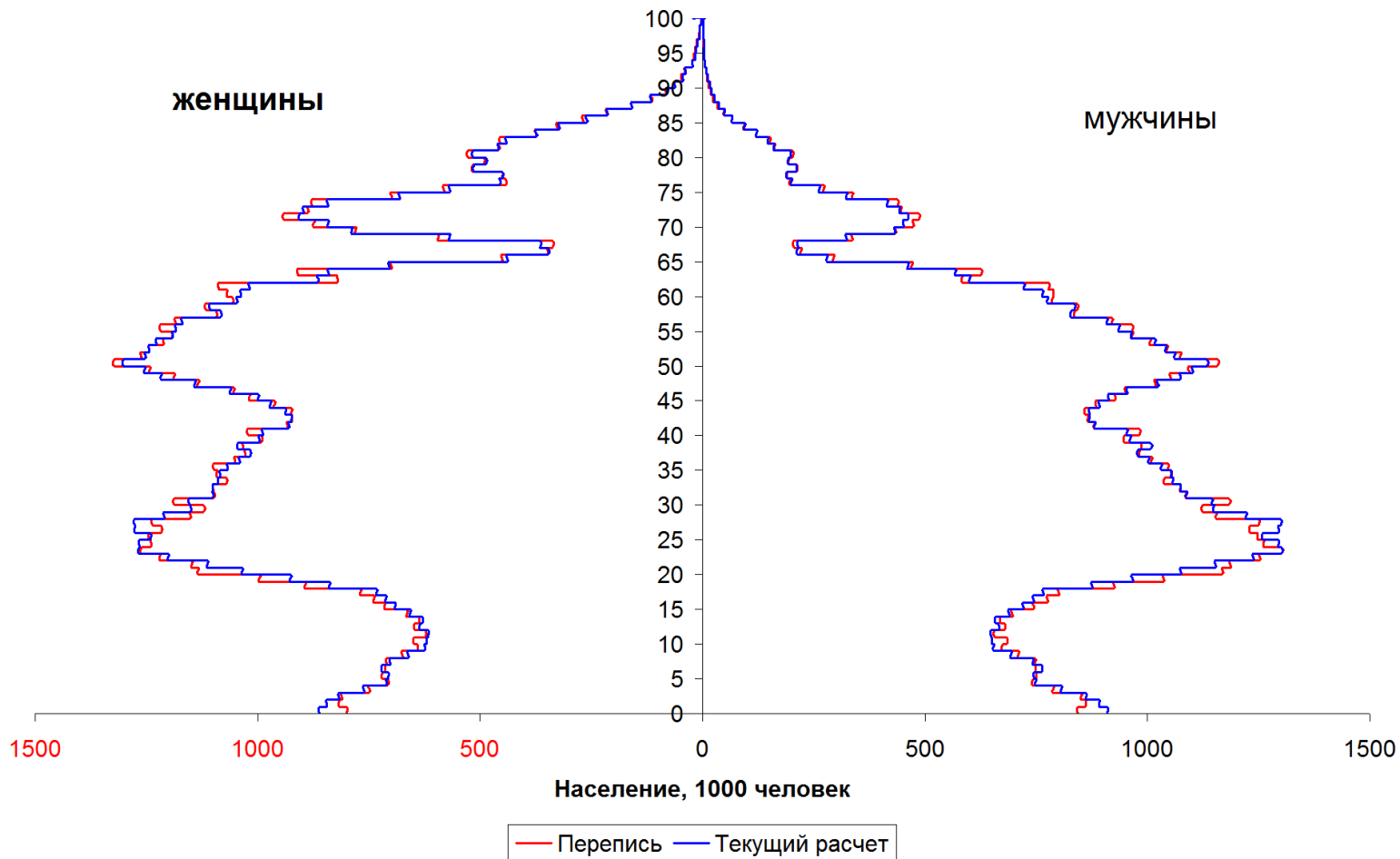
* Без учета данных по Чеченской республике о населении, так как в 1993-2002 годах данные о движении населения республики не собирались.

Сравнение итогов переписи населения 2002 г. и результатов текущей оценки населения России* на момент переписи

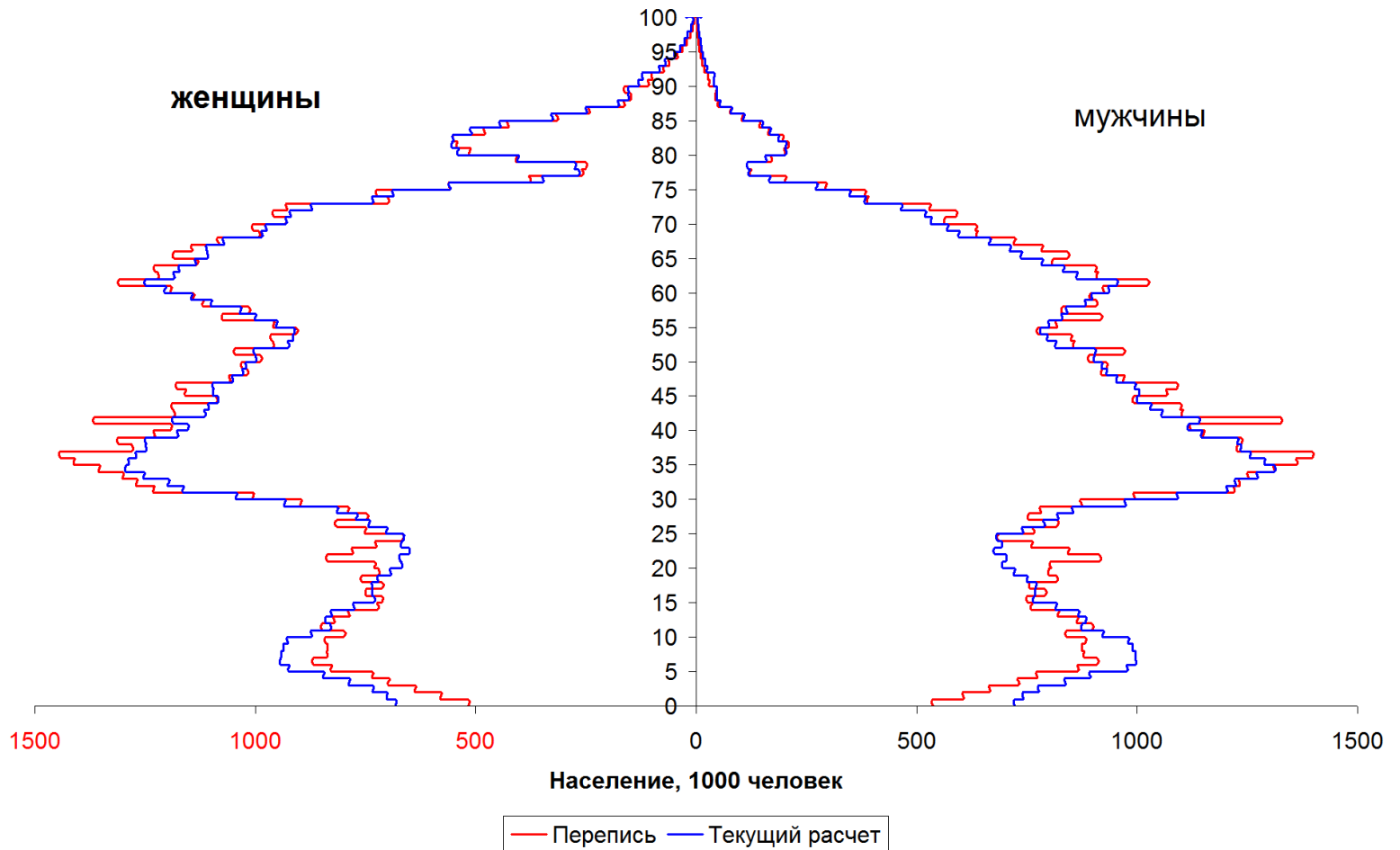


* Без учета данных по Чеченской республике, о населении, так как в 1993-2002 годах данные о движении населения республики не собирались.

Сравнение итогов переписи населения 2010 г. и результатов текущей оценки населения России на момент переписи



Сравнение итогов переписи населения 2021 г. и результатов текущей оценки населения России на момент переписи



Доля населения с неизвестным уровнем образования и брачным статусом, в процентах в трех переписях

Год переписи	Уровень образования в возрасте 10 лет и более	Брачный статус в возрасте 16 лет и более
2002	1,1	2,2
2010	2,9	3,1
2021	16,9	15,4

Федеральный закон от 25 января 2002 г. № 8-ФЗ «О Всероссийской переписи населения» с изменением от 28.11.2009 (Федеральный закон № 293-ФЗ) позволил проводить перепись по данным административных источников, но брать из них только пол и дату рождения.

Это привело к росту доли лиц с неизвестным образованием и брачным статусом.

**5 территорий с наивысшей долей населения в
возрасте 6 лет и старше с неизвестным уровнем
образования и 16 лет и старше с неизвестным
брачным статусом**

		Уровень образования неизвестен	Брачный статус неизвестен	
	г. Севастополь	37,7	29,6	
	г. Москва	34,2	32,8	
	Ханты-Мансийский	30,5	26,5	
	Приморский край	30,3	26,4	
	г. Санкт-Петербург	29,0	26,3	

Что определяет высокую долю населения с неизвестным образованием и брачным статусом (1)

Коэффициенты корреляции Кендала между

- (1) Доля лиц с неизвестным уровнем образования
- (2) Доля лиц с неизвестным брачным статусом
- (3) Стандартизованный коэффициент смертности от COVID-19

	(1)	(2)	(3)
(1)	1	0,88	0,26
(2)	0,88	1	0,23
(3)	0,26	0,23	1

Все показатели значимы на уровне не более 0,01 (двусторонний критерий)

Несколько слов о корреляции Кендалла

Если у нас есть n объектов, то из них можно выбрать $\frac{1}{2}n(n-1)$ разных пар объектов. Из 85 территорий можно сформировать 3570 пар территорий.

У нас есть две случайные величины A_i и B_i , i – номер территории.

Посчитаем число таких пар, что $A_i > A_j$ и $B_i > B_j$ или $A_i < A_j$ и $B_i < B_j$ (различия однонаправлены), и где это не так (различия разнонаправлены). Если $A_i = A_j$ или $B_i = B_j$ включим в каждое число по $1/2$.

Коэффициент корреляции Кендалла равен доле пар с однонаправленными различиями минус доля пар с разнонаправленными различиями среди всех пар.

Коэффициент 0,88 означает, что в 3352 парах (94% всех пар) доли переписанных с неизвестным образованием и с неизвестным брачным статусом различаются согласованно.

Коэффициент 0,23 означает, что в 2196 парах доли переписанных с неизвестным брачным статусом и СКС от COVID-19 различаются согласованно. Таких пар в два раза больше, чем тех, где различия не согласованы.

Что определяет высокую долю населения с неизвестным образованием и брачным статусом (2)

Коэффициенты корреляции Кендала между

- 1) Доля лиц с неизвестным уровнем образования в 2021 г.
- (2) Доля лиц с неизвестным брачным статусом в 2021 г.
- (3) Доля лиц с неизвестным уровнем образования в 2010 г.
- (4) Доля лиц с неизвестным брачным статусом в 2010 г.

	(1)	(2)	(3)	(4)
(1)	1	0,88	0,33	0,34
(2)	0,88	1	0,37	0,37
(3)	0,33	0,37	1	0,95
(4)	0,34	0,37	0,95	1

Все показатели значимы на уровне не более 0,01 (двусторонний критерий)

О ситуации в регионах России

Хотя численность населения по переписи 2021 г. оказалась существенно больше, чем по текущей оценке, в 46 из 85 территорий она была ниже. Перепись 2010 г. тоже дала общую численность населения России большую, чем по текущим расчетам. Из 83 непересекающихся территорий, где прошла перепись, также в 46 население оказалось меньше, а в 37 больше расчетной. Отметим, что 30 из 46 территорий совпадают.

Число регионов, где

	Всего	0-14	15-79	80+
Перепись < оценки	46	78	24	70
Перепись > оценки	39	7	61	15

Именно возрастная группа 15 - 79 определила превышение переписной численности над расчетной. Коэффициент Кендалла между общим относительным превышением и превышением в возрасте 15 – 79 лет равен 0,76 значим на уровне 0,01.

Относительная разность между текущей оценкой на момент переписи и итогами переписи 2021 г. в 5 регионах с наибольшими и наименьшими значениями этого показателя для всех возрастов

		Всего	0-14	15-79	80+	
1	Республика Карелия	-12,5	-19,2	-11,4	-7,2	
2	Архангельская область без автономного округа	-9,1	-12,1	-8,4	-10,5	
3	Республика Коми	-8,8	-14,7	-7,7	0,4	
4	Мурманская область	-8,4	-9,9	-8,4	-0,4	
5	Ямало-Ненецкий автономный округ	-7,6	-11,2	-6,7	6,5	
	Российская Федерация	1,1	-10,5	3,8	-5,7	
73	г. Москва	3,0	-14,7	7,0	-16,2	
78	г. Санкт-Петербург	4,1	-14,1	8,1	-9,3	
81	Ленинградская область	4,9	-7,4	7,7	-10,7	
82	г. Севастополь	5,4	-22,7	11,0	-6,1	
83	Калужская область	5,8	1,7	7,2	-6,2	
84	Республика Адыгея	6,3	3,3	7,6	-6,1	
85	Московская область	9,5	-20,3	15,3	3,0	

Насколько реально пересмотреть динамику показателей 2011-2021 годов так, чтобы они приводили к переписи 2021 г.?

Изменение численности поколений между 01.10.2010 и 01.10.2021 по данным текущего учета движения населения и с учетом переписи населения 2021 г., тыс. человек. Расчет по территории охваченной переписью 2010 г.

Возраст на 1 октября 2021 г.	Мужчины			Женщины		
	Рост когорты			Рост когорты		
	по текущей оценке	с учетом переписи	различие	по текущей оценке	с учетом переписи	различие
0-4	0,8	-651,6	-652,4	5,2	-568,9	-574,1
5-9	-26,4	-546,9	-520,4	-15,3	-474	-458,7
10-14	10,7	-164,1	-174,8	14	-152,8	-166,7
15-19	21,5	161,4	139,9	39,5	72,2	32,7
20-24	29,6	587,7	558,1	72,7	479,3	406,5
25-29	162,6	-16,6	-179,2	99,8	135,4	35,6
30-34	64,8	-30,5	-95,3	98,7	303,6	204,9
35-39	-98,8	118,4	217,2	39,1	477,2	438
40-44	-253,1	32,4	285,5	-29,6	335,7	365,4
45-49	-342,2	-176,3	165,9	-79,6	64,6	144,2
50-54	-412,6	-274,7	137,9	-126,6	-17,8	108,8
55-59	-668,8	-555,3	113,5	-229	-154,4	74,6
60-64	-1076	-880	196	-435,1	-301,8	133,3
65-69	-1296,7	-965,8	330,9	-639,2	-487,1	152,1
70-74	-1327	-1126,3	200,8	-841,3	-745,2	96,1
75-79	-725,5	-655,7	69,8	-643,7	-645	-1,3
80-84	-1377,3	-1377,7	-0,4	-1795,8	-1880,7	-84,9
85+	-1853,6	-1940,3	-86,7	-4798,6	-4933,6	-134,9

Влияет ли доля переписанных по административным данным на относительную разность между текущей оценкой на момент переписи и итогами переписи 2021 г. в возрастах до 15 лет и 80 лет и старше (1)

Коэффициенты корреляции Кендала между

- (1) Доля лиц с неизвестным уровнем образования в 2021 г.
- (2) Доля лиц с неизвестным брачным статусом в 2021 г.
- (3) Доля лиц с неизвестным уровнем образования в 2010 г.
- (4) Доля лиц с неизвестным брачным статусом в 2010 г.
- (5) Относительная разность между текущей оценкой на момент переписи и итогами переписи 2021 г. в возрастах до 15 лет.
- (6) Относительная разность между текущей оценкой на момент переписи и итогами переписи 2021 г. в возрастах 80 лет и старше

(Знаменатель – средняя по двум источникам численность возрастной группы)

Влияет ли доля переписанных по административным данным на относительную разность между текущей оценкой на момент переписи и итогами переписи 2021 г. в возрастах до 15 лет и 80 лет и старше (2)

	(1)	(2)	(3)	(4)	(5)	(6)
(5)	0,19	0,21	☺	☺	1	☺
(6)	-0,23	-0,26	-0,23	-0,20	☺	1

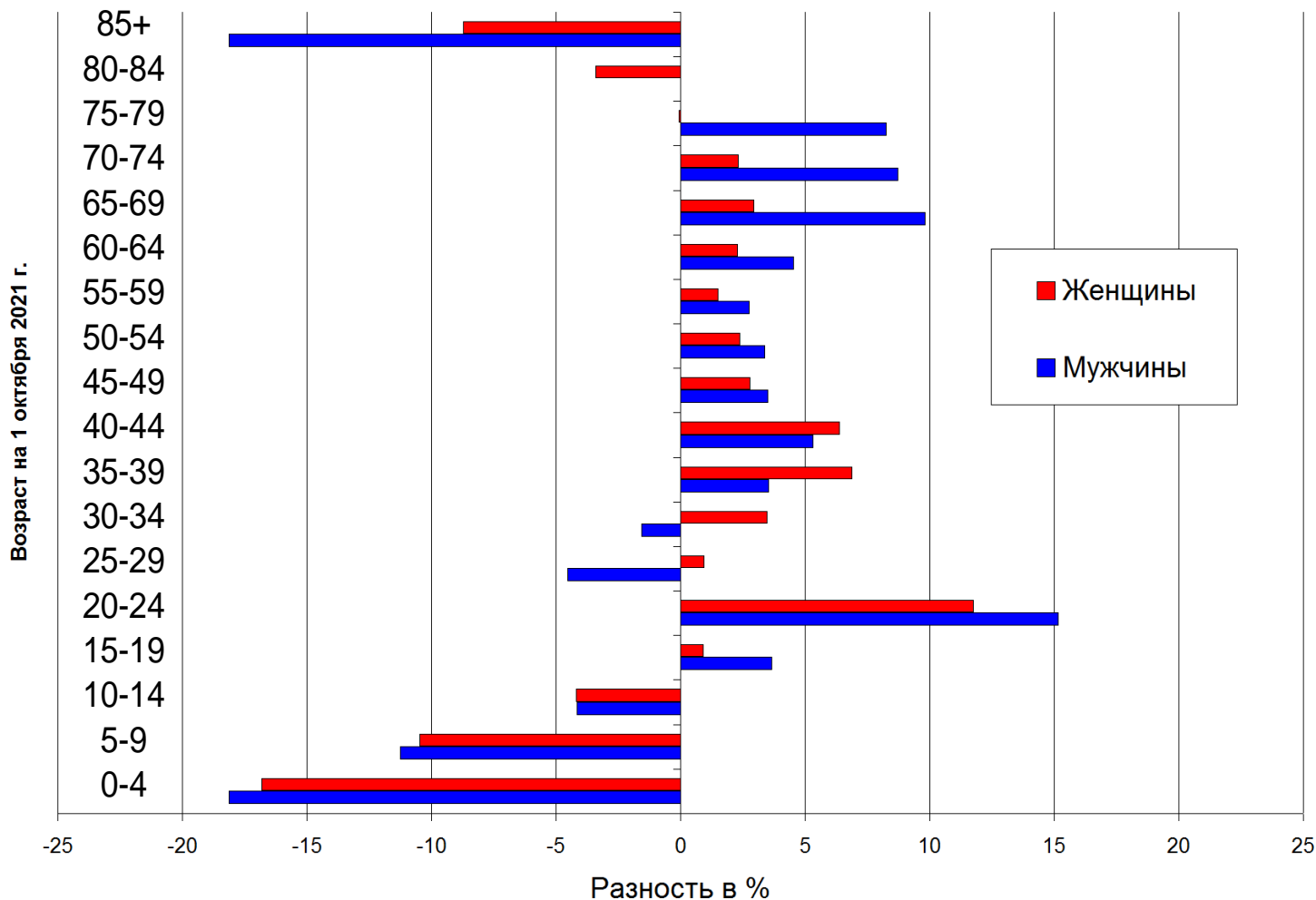
Указаны только показатели значимые на уровне не более 0,05 (двусторонний критерий), значимые на уровне не более 0,01 выделены полужирным шрифтом. Незначимые заменены на ☺.

Различия между результатами переписи и текущей оценкой по общей численности населения и по числу лиц в возрасте 15-79 лет с данными о числе лиц с неизвестным образованием или брачным статусом значимо не коррелирует.

Изменение численности поколений между 01.10.2010 и 01.10.2021 по данным текущего учета смертей и два варианта чистой миграции в Россию поколений, в период с 01.10.2010 по 30.09.2021, тыс. человек. Расчет по территории, охваченной переписью 2010 г.

Возраст на 1 октября 2021 г.	Мужчины			Женщины		
	Числа умерших	Чистая миграция		Числа умерших	Чистая миграция	
		Данные статистик и	С учетом переписи		Данные статистики миграции	С учетом переписи
0-4	18,1	18,9	-633,5	13,7	18,9	-555,2
5-9	40,8	14,4	-506,1	30,9	15,7	-443
10-14	22,6	33,3	-141,5	16,4	30,4	-136,3
15-19	16,3	37,8	177,7	9,4	48,8	81,5
20-24	31,3	60,9	619	13,4	86,2	492,7
25-29	71,6	234,2	55	23,7	123,5	159,1
30-34	182,8	247,6	152,3	57,4	156,1	361
35-39	306,8	208	425,2	101,9	141,1	579,1
40-44	411,5	158,4	443,9	141,6	111,9	477,3
45-49	460	117,8	283,7	168	88,4	232,6
50-54	502,8	90,2	228,1	196	69,4	178,2
55-59	739,7	70,9	184,4	301	71,9	146,6
60-64	1132,8	56,9	252,8	508,8	73,8	207
65-69	1331,3	34,6	365,5	693,9	54,7	206,7
70-74	1345,4	18,4	219,2	880,1	38,8	134,9
75-79	732,8	7,3	77,1	662,1	18,4	17,1
80-84	1382,7	5,4	5	1817	21,2	-63,7
85+	1865,7	12,1	-74,6	4816,4	17,8	-117,1

Разность оценки миграционного прироста численности когорты за межпереписной период по данным статистики миграции минус с учетом переписи 2021 г. в процентах к численности когорты на конец периода (взята средняя по двум источникам численность)



Некоторые итоги

Обнаруженная связь между долей переписанных по административным данным и расхождением итогов переписи с текущей оценкой в детских возрастах позволяет считать, что расхождение есть следствие неполного учета детей в административных базах данных. Иначе, если дети были включены в базу данных, но выехали из страны, то это вошло в статистику миграции. Представляется, что правильнее всего было бы в возрастах до 15 лет ориентироваться на результаты текущего расчета.

Коррекция текущего расчета в когортах, которым при переписи было от 15 до 79 лет – обычная задача, которая решается пропорциональным распределением добавочных мигрантов. Но смущают значительные между возрастными перепадами. Возможно, следует применить выравнивание скачков.

Сложнее всего с пожилыми возрастами. Неоднократно высказывалось предположение, что перепись 2010 г. зависила число пожилых. Наиболее вероятно, что перепись 2021 г. подтвердила это предположение. Вероятно, были уточнены административные источники, а пустующие квартиры получили новых хозяев. Но пересмотр итогов переписи 2010 г. вряд ли возможен. Рациональнее пересчитать оценки за 2011-2020 гг. и смириться с неплавностью ряда между 2010 и 2011 годами.

Полезный совет

Высокая доля населения с неизвестным уровнем образования, брачным статусом и т.д., женщин с неизвестным числом рожденных детей заставляет особенно аккуратно относиться к распределению неизвестных. Дело в том, что образование или число детей и доля неизвестных не независимые переменные. Например, образование выше, а число рождений ниже там, где выше доля неизвестных. Поэтому для целей анализа правильнее всего сначала распределить неизвестных по самым мелким из доступных территорий, например, по городскому и сельскому населению субъектов федерации, и только потом суммировать по региону в целом, по группам регионов и по стране в целом.

Перепись 2021 г. проведена в сложных условиях пандемии, что не могло не сказаться на качестве результатов. Однако, на наш взгляд, высказывания о недостоверности всех итогов переписи ничем не обоснованы. Итоги переписи не идеальны, но в целом точность переписи 2021 г. не ниже, чем переписей 2002 и 2010 годов.

Спасибо за внимание!