

НАСКОЛЬКО СЕРЬЕЗНОЕ ВЛИЯНИЕ ОКАЗЫВАЮТ МЕЖИССЛЕДОВАТЕЛЬСКИЕ РАСХОЖДЕНИЯ НА РЕЗУЛЬТАТЫ КРАНИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ? (НЕКОТОРЫЕ ИТОГИ СЕМИНАРА ПО КОННЕКСИИ КРАНИОМЕТРИЧЕСКИХ ПРИЗНАКОВ В МАЭ РАН)

И.Г. Широбоков

*Музей антропологии и этнографии им. Петра Великого (Кунсткамера) РАН,
Санкт-Петербург, Россия*

В 2015 году в стенах Музея антропологии и этнографии им. Петра Великого (Кунсткамера) РАН прошел научно-практический семинар, посвященный оценке межисследовательских расхождений в краниометрии. Участники семинара измеряли две краниологические серии русских и бурят, каждая из которых включала в себя 10 черепов взрослых человек. Программа измерений включала в себя признаки, наиболее часто используемые краниологами (№ по Мартину и др.: 1, 8, 17, 9, 11, 45, 48, 51, 52, 55, 54, SC, SS, MC, MS, 72, 73, 75(1)). Результаты сопоставления среднего относительного размаха вариации признаков показали, что в наибольшей степени между исследователями варьируют признаки, характеризующие степень выступления носовых костей и переносья, а также вертикальные углы лицевого скелета, в меньшей степени – форма орбит и грушевидного отверстия. Наибольшей надежностью отличаются измерения основных диаметров и углы горизонтальной профилировки. По всей вероятности, на расхождения признаков повлияли как случайные погрешности в измерении признаков малой величины, так и методические расхождения, связанные с трудностями локализации некоторых краниометрических точек, в частности, maxillofrontale, ektokonchion и nasospinale. При этом фактор невнимательности при измерении и вводе данных вносит не меньший вклад в расхождения между исследователями, чем систематические методические расхождения. Влияние межисследовательских расхождений на результаты внутригруппового анализа оценивалось при помощи метода главных компонент, на результаты межгруппового – канонического дискриминантного анализа. Результаты оценки однородности выборки, полученные разными исследователями, вероятно, в большинстве случаев могут считаться сопоставимыми между собой при отсутствии различий в определении половой принадлежности черепов. На результаты межгрупповых сопоставлений весьма значительный эффект оказывает суммарное влияние эффекта выборки и межисследовательских расхождений в измерениях. Некорректно судить о сходстве конкретных популяций на основании сближения соответствующих выборок в пространстве первых канонических векторов. Сокращение краниометрической программы за счет исключения из нее наименее надежных признаков не увеличивает достоверность результатов, поскольку в этом случае возрастает влияние расхождений между исследователями в определении половой принадлежности черепов. Решение этих проблем может заключаться во введении в качестве обязательной исследовательской практики повторного измерения всех признаков в каждой изучаемой серии, а также использовании сравнительных данных лишь тех исследователей, сопоставимость измерений которых не вызывает сомнений.

Ключевые слова: краниометрия, межисследовательская сопоставимость, внутригрупповой анализ, межгрупповой анализ

Введение

В октябре 2015 года в стенах Музея антропологии и этнографии им. Петра Великого (Кунсткамера) (МАЭ) РАН состоялась Всероссийская научно-практическая конференция «Палеоантропологические и биоархеологические исследования: традиции и новые методики» (VI Алексеевские чтения). В рамках конференции был организован научно-практический семинар, посвященный оценке межисследовательских расхождений в современной краниометрии. Единственный за многие десятилетия семинар по коннекции краниометрических программ проходил в 2013 г. в стенах Института этнологии и антропологии РАН [Пежемский, Харламова 2013]. Его участники особое внимание уделили обсуждению техники измерения, в частности, методическим расхождениям в определении конкретных краниометрических точек. Главная цель семинара 2015 г. заключалась в статистической оценке надежности наиболее часто используемых в исследованиях признаков, а также оценке влияния межисследовательских расхождений на результаты внутригруппового и межгруппового анализов.

Участники семинара измеряли две краниологические серии (численностью по 10 черепов), находящиеся на постоянном хранении в МАЭ РАН: серию русских Старой Ладogi (№ 5720-2, 10, 42, 47, 50, 55, 59, 68, 69, 75) и сборную серию бурят (№ 4238-1, 5; 6310-8, 12, 18, 53, 69, 78, 81, 91). Для измерения были предложены следующие признаки (здесь и далее № по Мартину и др.): 1, 8, 17, 9, 11, 45, 48, 51, 52, 55, 54, SC, SS, MC, MS, 72, 73, 75(1).

В работе семинара приняли участие исследователи из Барнаула, Казани, Москвы, Пензы, Санкт-Петербурга, Самары, Севастополя и Тюмени. Всего в измерении серии русских приняли участие 13 человек, бурятской серии – 9 человек. При проведении анализа вновь полученные данные измерений были дополнены неопубликованными индивидуальными измерениями В.П. Алексеева и Н.Н. Мамоновой, хранящимися в архиве Отдела антропологии МАЭ РАН.¹

Измерения черепов осуществлялись участниками независимо друг от друга, а их результаты

фиксируются на индивидуальных бланках, впоследствии переведенных в электронный формат. Перед проведением статистического анализа из данных были исключены все измерения признаков, ошибочность которых была очевидной даже при отсутствии возможности сопоставления размеров. Кроме того, была проведена корректировка измерений, ошибочность которых объяснялась невнимательностью участников при определении значений на толстотном циркуле (смещение реального значения на 5 мм). Введение последней поправки было обусловлено тем обстоятельством, что в процессе работы несколько участников отметили, что некоторые инструменты отличаются от тех, которыми они обычно пользуются. Корректировка проводилась следующим образом: если разница в измерениях какого-то признака у одного из исследователей составляла около 5 мм относительно измерений большинства коллег, полученное им значение признака изменялось в сторону среднего значения ровно на 5 мм. Целесообразность введения поправки может быть проиллюстрирована следующим примером: один из участников семинара, допустивший максимальное число явных искажений при измерении первой серии черепов, при исследовании второй серии явных ошибок уже не допускал.

Вместе с тем, поскольку невнимательность при работе с инструментом может играть важную роль и в обычной исследовательской практике, при оценке расхождений между конкретными участниками семинара наличие (“1”) или отсутствие (“0”) очевидных ошибок измерения фиксировалось в качестве самостоятельного признака.

С учетом проведенной корректировки расхождения в результатах измерений теоретически могут быть объяснены следующими факторами:

1. Различия в определении краниометрических точек (методические расхождения).
2. Ошибки инструмента (плохая верификация)².
3. Определение приблизительных значений признаков в случаях неудовлетворительной сохранности отдельных костей или отказ от их измерения.

¹В статье результаты измерений приводятся под условными номерами без указания авторов, поскольку у читателей, не владеющих всей исходной информацией о проведенных измерениях, может сложиться превратное впечатление о целесообразности использования публикуемых некоторыми участниками семинара данных в собственных исследованиях.

²При подготовке к семинару проводилась верификация всех предназначенных к работе наборов краниометрических инструментов. Тем не менее, в ходе работы выяснилось, что при использовании некоторых циркулей все же возникает небольшая погрешность при измерении. Таким образом, условия работы были неумышленно приближены к реальной ситуации, когда мы используем литературные данные, опубликованные другими авторами, и не можем быть уверены в том, что они проводили верификацию инструментов перед изучением каждой серии.

Таблица 1. Средние значения признаков выборки русских Старой Ладogi по данным аналогичных измерений 14 исследователей

Признаки	V_R	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1,2	180,4	180,9	180,5	180,4	181,1	180,4	180,5	180,4	180,6	180,0	180,5	180,4	180,4	180,5
8	1,9	145,5	144,7	143,7	144,8	145,8	144,0	144,7	145,5	144,3	144,8	145,4	144,5	144,4	144,7
17	1,0	136,8	137,1	136,8	136,3	136,6	136,3	136,4	136,4	136,8	136,3	136,2	136,2	136,8	136,5
9	2,3	102,9	102,5	102,4	102,3	102,8	102,5	102,3	102,4	102,6	102,2	102,2	102,5	101,2	102,3
11	4,4	124,9	125,1	126,4	126,0	125,7	125,2	125,4	125,7	125,7	125,4	125,4	124,5	123,9	126,1
45	1,5	135,2	135,3	135,8	135,2	135,3	134,2	135,0	135,3	135,2	135,1	134,9	135,1	135,3	135,2
48	4,0	73,1	72,8	73,4	73,0	73,2	71,9	72,9	72,8	73,5	72,8	72,7	73,5	73,3	73,1
51	9,1	42,8	44,4	43,9	43,6	43,0	43,9	43,8	43,9	44,2	43,7	42,8	44,1	44,1	45,2
52	10,0	32,2	33,3	32,2	32,4	32,2	32,9	32,6	32,8	32,3	32,5	32,0	31,1	32,6	32,5
55	6,8	51,0	51,6	51,3	50,8	51,6	51,5	50,5	50,8	51,2	51,1	50,6	51,0	50,5	51,4
54	4,8	24,2	24,0	24,2	24,4	24,2	24,1	24,3	24,5	24,3	24,2	24,6	24,0	24,4	24,4
43(1)	1,9	102,8	102,0	101,5	101,7	102,0	101,7	101,1	102,1	101,9	101,9	102,0	102,1	101,8	102,0
sub.n/43(1)	12,4	19,6	19,2	19,8	19,7	19,4	18,8	19,6	19,8	19,4	19,8	19,6	19,4	19,7	19,7
zm'-zm'	3,9	98,0	98,2	98,0	98,4	98,5	98,8	97,9	98,3	98,1	98,6	98,9	98,2	97,0	98,0
sub.ss/zm'-zm'	13,7	24,2	24,1	23,8	24,1	24,6	23,2	25,1	23,8	23,9	24,1	24,5	23,8	24,5	24,4
SC	16,9	10,6	10,6	10,1	11,1	10,8	11,4	10,5	10,8	10,7	10,9	10,9	10,8	10,2	10,7
SS	29,3	5,4	5,4	4,8	5,3	5,5	5,4	5,3	5,2	5,3	5,3	5,3	5,4	5,2	5,4
MC	14,5	20,8	19,4	21,7	21,5	21,8	21,8	21,3	20,8	20,8	21,2	22,1	–	20,4	20,3
MS	51,6	–	8,2	9,0	8,5	9,2	11,4	9,0	8,2	8,3	8,6	9,0	–	8,2	8,1
72	6,5	87,0	86,4	–	84,9	–	83,9	84,5	85,1	84,3	85,7	85,7	85,7	84,5	84,7
73	7,0	86,6	88,3	–	90,0	–	84,7	89,2	87,1	89,0	89,4	88,2	87,9	87,2	86,4
75	9,8	50,6	52,4	50,6	48,6	–	50,3	49,4	49,4	49,2	50,2	51,0	51,0	49,4	48,4
75(1)	12,5	33,5	35,5	–	35,5	32,3	36,0	34,0	34,0	34,5	35,0	35,0	33,0	35,0	35,5
77	3,3	138,3	138,8	137,4	137,7	138,4	139,6	137,6	137,6	138,3	137,6	137,9	138,5	137,8	137,9
zm	4,8	127,6	127,9	128,3	128,0	127,1	129,8	125,8	128,5	128,3	128,0	127,5	128,5	126,8	127,3
SS:SC	21,6	51,4	51,2	47,4	48,6	51,0	47,3	51,0	48,3	50,0	49,1	49,0	50,2	51,3	50,4
MS:MC	34,9	–	42,8	42,7	40,4	43,2	49,8	43,0	40,6	40,9	41,3	42,0	–	41,1	40,7
Err	–	нет	да	нет	нет	нет	да	да	нет	да	нет	нет	да	да	нет

Примечания. V_R – усредненный относительный размах вариаций признака (%). Измерения конкретных исследователей обозначены условными номерами от 1 до 14. Серой заливкой отмечены крайние значения каждого признака.

- Неявные ошибки измерений или ввода данных, связанные с невнимательностью участников. При сопоставлении средневыворочных значений признаков этот список должен быть также расширен за счет включения дополнительного фактора:
- Расхождения в определении половой принадлежности черепов.

По всей вероятности, отдельного рассмотрения заслуживает также вопрос о сопоставимости измерений, проведенных одним исследователем в разное время. Но поскольку в рамках одного семинара осуществить оценку внутриисследовательской ошибки невозможно, эта задача должна

быть поставлена при организации следующих семинаров.

Общая оценка расхождений между исследователями

На первом этапе исследования проводилась общая оценка надежности измерения признаков путем сопоставления их среднегрупповых значений. Поскольку оценивалось влияние именно измерительной техники, различия между исследователями в определении половой принадлеж-

Таблица 2. Средние значения признаков выборки бурят по данным аналогичных измерений 10 исследователей

Признаки	V_R	1	2	3	4	5	6	7	8	9	10
1	1,5	182,0	183,0	182,6	182,9	183,1	181,8	182,6	182,6	182,8	182,4
8	1,9	152,9	153,8	154,4	154,1	154,5	153,1	154,4	154,3	153,9	154,0
17	1,7	132,4	132,8	132,6	132,2	132,0	131,5	132,5	132,0	132,4	132,5
9	1,8	100,1	99,9	100,9	99,8	99,4	99,6	100,1	99,9	100,1	99,9
11	3,9	139,5	142,3	140,9	138,3	139,3	138,5	142,6	140,9	138,4	140,9
45	1,0	145,4	145,6	145,7	145,6	145,3	144,6	145,5	145,3	145,4	145,6
48	4,1	78,8	79,1	79,5	78,8	78,0	76,6	79,0	78,8	79,2	79,5
51	8,5	45,4	43,6	44,1	44,5	44,1	43,4	44,7	44,6	43,1	45,1
52	7,5	36,1	36,6	36,2	36,2	36,6	35,6	35,7	35,0	35,6	36,1
55	8,1	56,1	57,1	56,1	56,0	57,3	56,7	54,7	56,2	58,5	56,3
54	4,9	27,4	27,8	27,7	28,0	27,6	28,2	27,7	28,1	27,9	27,6
43(1)	2,9	102,9	102,7	102,4	102,4	101,1	102,3	–	102,9	103,1	102,7
sub.n/43(1)	16,1	15,4	15,9	15,6	15,7	15,6	15,8	–	15,6	14,1	15,4
zm'-zm'	6,2	106,1	104,0	106,3	106,6	104,8	105,4	–	107,1	106,9	106,4
sub.ss/zm'-zm'	26,3	19,1	21,9	19,0	18,4	19,0	19,2	–	18,2	18,5	19,6
SC	26,6	9,1	9,9	8,7	8,9	8,8	9,3	–	9,2	9,8	9,1
SS	42,4	2,7	2,7	2,6	2,5	2,7	2,3	–	2,7	3,0	2,8
MC	25,3	17,9	20,4	21,7	20,4	21,0	21,6	–	21,1	–	20,6
MS	68,3	4,1	6,0	6,2	5,3	6,8	5,4	–	5,6	–	5,4
72	4,8	87,5	89,3	–	87,5	86,2	87,8	87,7	88,0	87,8	86,7
73	7,6	90,5	88,4	–	92,1	90,4	93,1	93,9	91,6	91,4	90,4
75	5,7	65,8	–	–	65,5	64,8	67,8	64,8	66,3	65,8	65,0
75(1)	22,5	23,3	23,5	20,3	23,5	21,0	22,3	24,8	23,7	23,8	23,8
77	3,6	146,6	145,7	146,1	145,9	145,8	145,7	–	146,3	149,6	146,7
zm	7,7	140,4	134,2	140,1	141,8	139,9	139,9	–	142,3	141,9	139,4
SS:SC	40,0	30,9	28,2	30,2	28,1	31,3	26,1	–	29,1	31,0	31,1
MS:MC	65,4	23,0	30,2	28,8	25,8	32,4	25,0	–	26,3	–	26,3
Err	–	нет	да	нет	нет	нет	да	да	да	да	нет

Примечания. V_R – усредненный относительный размах вариаций признака (%). Измерения конкретных исследователей обозначены условными номерами от 1 до 14. Серой заливкой отмечены крайние значения каждого признака.

ти черепов не учитывались. Кроме того, в случае, если кто-то из участников считал недопустимым проводить измерение какого-то признака на конкретном черепе, из сравнительного анализа исключались все значения данного признака, полученные другими участниками. Средние значения признаков в сериях русских и бурят, полученные с учетом перечисленных выше поправок представлены в табл. 1 и 2.

У каждого из участников семинара средневыборочные значения отдельных признаков входят в число минимальных или максимальных в рамках группы, что само по себе не является негативным результатом. С позиций теории вероятности этот результат является тем более ожидаемым, чем больше признаков включает в себя краниометрическая программа и чем уже круг исследователей,

результаты измерений которых сравниваются. Значительно больше настораживает тот факт, что у пяти участников, измерявших черепа русских, и шести участников, исследовавших бурятскую серию, присутствуют явные ошибки измерений, вынужденно подвергнутые корректировке перед проведением расчетов.

Сопоставление усредненных показателей относительного размаха вариаций признаков показывает, что наибольшие расхождения между исследователями возникают при измерении максиллофронтальной ширины и высоты, угла выступания носа к линии профиля, симметрических ширины и высоты, вертикальных лицевых углов, а также высот, используемых для вычисления зигомаксиллярного и назомаллярного углов (рис. 1). Примечательно, что за счет высокой надежности измерения

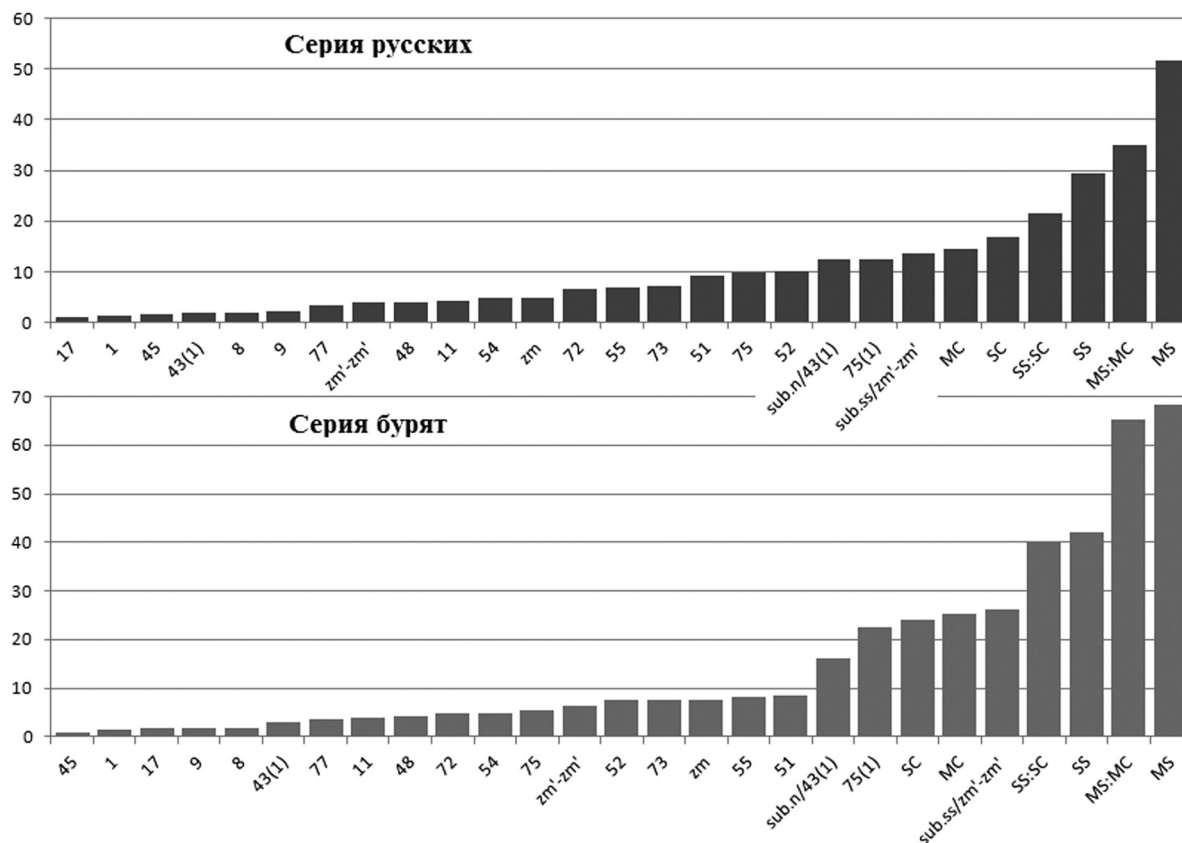


Рис. 1. Усредненный относительный размах вариации краниометрических признаков, измеренных разными исследователями (%)

соответствующих хорд последнее обстоятельство не сказывается существенным образом на вариативности измерений самих углов горизонтальной профилировки. Наименее вариативными являются значения основных диаметров, ширины лба и широтные размеры лицевого скелета. Исключение составляет ширина орбиты, расхождения в измерениях которой между исследователями довольно значительны, особенно в выборке русских ($V_R=10.0\%$). Высота носа ожидаемо, хотя и незначительно, сильнее колеблется в выборке бурят ($V_R=8.1$ и 6.8% соответственно).

Относительная вариативность измерений возрастает с уменьшением расстояния между краниометрическими точками. Расхождения в 1 мм для продольного диаметра и симотической ширины имеют разное значение и на практике. При проведении межгруппового анализа методами многомерной статистики нормирование величин исходных признаков значительно увеличивает значение погрешности измерения именно признаков с невысокой размерностью. При этом какие-либо систематические расхождения в измерениях, которые позволили бы разделить исследователей

на отдельные группы (например, обусловленные различиями в определении краниометрических точек или использованием разных наборов инструментов, некоторые из которых не прошли верификацию), отсутствуют. Об этом свидетельствуют результаты анализа, проведенного при помощи метода главных компонент (рис. 2).

Вопреки ожиданиям, в обоих случаях – как при анализе выборки русских, так и бурят – изменчивость измерений в пространстве первой компоненты преимущественно определяется не наименее, а наиболее надежными признаками, т.е. теми признаками, в измерении которых большинство исследователей допускает незначительную погрешность. По всей вероятности, эти результаты обусловлены отклонениями в измерениях единичных участников, не поддающимися строгому объяснению. Действительно, если различия в размерах верхней высоты лица можно связать с тем, что некоторые исследователи измеряют расстояние не от нижней, а от передней точки альвеолярного края, то объяснить отклонения в значениях скулового и высотного диаметров с методической точки зрения представляется затруднительным.

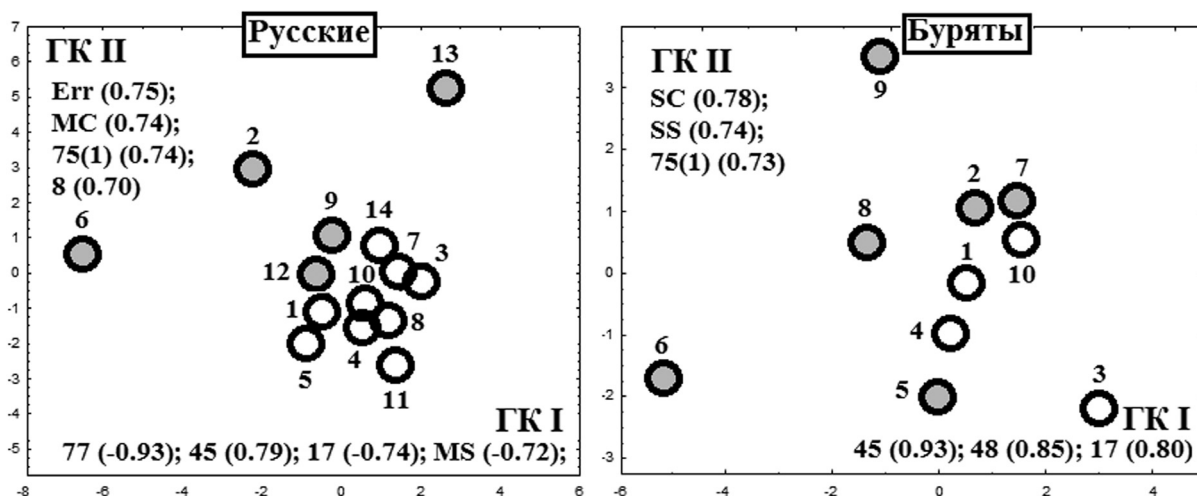


Рис. 2. Различия между исследователями по результатам измерения выборки русских и бурят по результатам анализа главных компонент

Примечания. Err – наличие явных ошибок измерения; серой заливкой отмечены измерения тех исследователей, которые допускали явные ошибки при измерении. Измерения конкретных исследователей обозначены условными номерами, которые не совпадают на графиках.

Вторая компонента у русских в наибольшей степени коррелирует с наличием явных ошибок измерения, а также значениями поперечного диаметра, угла выступания носа и максиллофронтальной ширины, у бурят – с величинами симметрических ширины и высоты, а также угла выступания носа. Таким образом, во второй компоненте на первый план выдвинулись признаки с наиболее высокими показателями относительного размаха вариаций измерений. Включение в эту группу признаков поперечного диаметра может быть связано как со случайными отклонениями, так и с завышением признака при определении точек *europ* на отклонившейся наружу чешуе височных костей.

Исключение из анализа признака «наличие/отсутствие явных ошибок измерения» не меняет существенным образом картину расхождений между исследователями. Обособленное положение на графике наборов измерений тех участников семинара, которые допускали явные ошибки измерений, позволяет предположить, что те же краниологи одновременно оказались более склонны делать случайные ошибки, не поддающиеся корректировке. Следовательно, наибольший вклад в межисследовательские расхождения вносят не различия в определении краниометрических точек, а невнимательность отдельных исследователей.

Внутригрупповой анализ

Оценка межисследовательских расхождений на внутригрупповом уровне сопоставления в первую очередь призвана выяснить являются ли сопоставимыми заключения исследователей, использующих одинаковые статистические методы, относительно однородности выборок. С этой целью матрицы индивидуальных значений признаков каждого из участников семинара были подвергнуты отдельной статистической обработке при помощи анализа главных компонент. Всего было проведено 14 вариантов анализа серии русских и 9 вариантов анализа серии бурят.³

В рамках каждого анализа при отсутствии индивидуального измерения применялась подстановка среднего значения. Использовались следующие признаки: 1, 8, 17, 9, 45, 48, 55, 54, 51, 52, SS, SC, 77, zm, 75(1). Анализ проводился без учета расхождений в определении пола. Теоретически, именно этот фактор может оказывать существенное влияние на расхождения в оценках однородности выборки. К сожалению, небольшая численность

³Данные одного из исследователей пришлось исключить из анализа, поскольку он по каким-то причинам не фиксировал значительную часть предложенных для измерения признаков (SS, SC, MC, MS, 77, zm). В другом случае исследователь не фиксировал признак 75(1), поэтому при анализе использовалась подстановка средних значений.

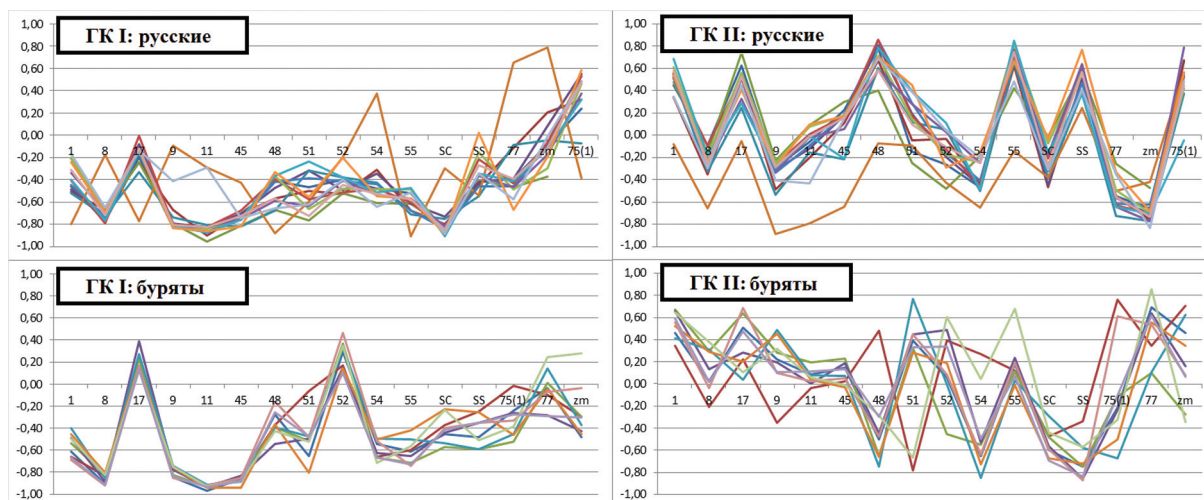


Рис. 3. Разброс значений коэффициентов корреляции между исходными признаками и первыми двумя главными компонентами

Примечания. Кривая каждого цвета соответствует нагрузкам, полученным по результатам анализа данных одного исследователя.

измеренных серий не позволяет рассмотреть влияние этих расхождений на результаты.

Оценка однородности выборки проводилась путем тестирования нормальности распределения значений первых двух главных компонент при помощи критерия Шапиро-Уилка. В 13 из 14 вариантов анализа серии русских, и в 8 из 9 вариантов анализа бурятских черепов распределение новых признаков не отличается от нормального вида. В одном случае было выявлено статистически значимое отклонение от нормального вида для значений первой ГК ($p < 0.03$), и в одном – для второй ($p < 0.03$). Таким образом, два исследователя теоретически имели основания заподозрить, что изученная ими выборка неоднородна, хотя низкий уровень значимости и небольшая численность черепов не позволяли бы им говорить об этом с уверенностью.

Сопоставление различных вариантов анализа данных, полученных по результатам измерений одной и той же серии разными исследователями, является также важным для оценки перспективности изучения выборок при помощи еще не вошедших в общую практику методов. В частности, в изучении вариаций элементов факторной структуры и возможностей установления стандартов колебаний факторных нагрузок в однородных и смешанных группах при заданном наборе признаков [Козинцев, 2016].

На рисунке 3 в виде кривых представлены результаты наложения элементов факторной структуры первых двух главных компонент для всех рассчитанных вариантов анализа серий русских и бу-

рят соответственно. На первый взгляд наибольшие колебания в нагрузках между исследователями приходится на признаки лицевого скелета, тогда как основные размеры мозгового отдела связаны с главными компонентами относительно стабильными корреляциями. Однако в действительности вариативность нагрузок, связанных с размерами мозгового отдела, уступает остальным только в вариантах анализа бурятских черепов. Основные же различия заключаются в том, что колебания нагрузок, связанных с основными диаметрами черепа, происходят более согласованно между собой, чем корреляции отдельных признаков лицевого скелета. Из-за небольшой численности исследуемых выборок различия между коэффициентами корреляции, полученными для аналогичных пар признаков, в большинстве случаев не достигают порога статистической значимости. В случае, если при работе исследователей с более крупными выборками нагрузки варьируют сходным образом (о чем мы можем судить лишь гипотетически), различия коэффициентов будут вполне реальными.

Полученные результаты свидетельствуют о некоторых сложностях в решении вопроса об установлении стандартов колебаний факторных нагрузок в однородных группах. По всей видимости, выработка таких стандартов должна опираться на данные измерений одного исследователя или ограничиваться кругом признаков с невысокой долей межисследовательских расхождений. В случае, если такие стандарты будут когда-нибудь выработаны, необходимо также продумать, каким образом учитывать расхождения

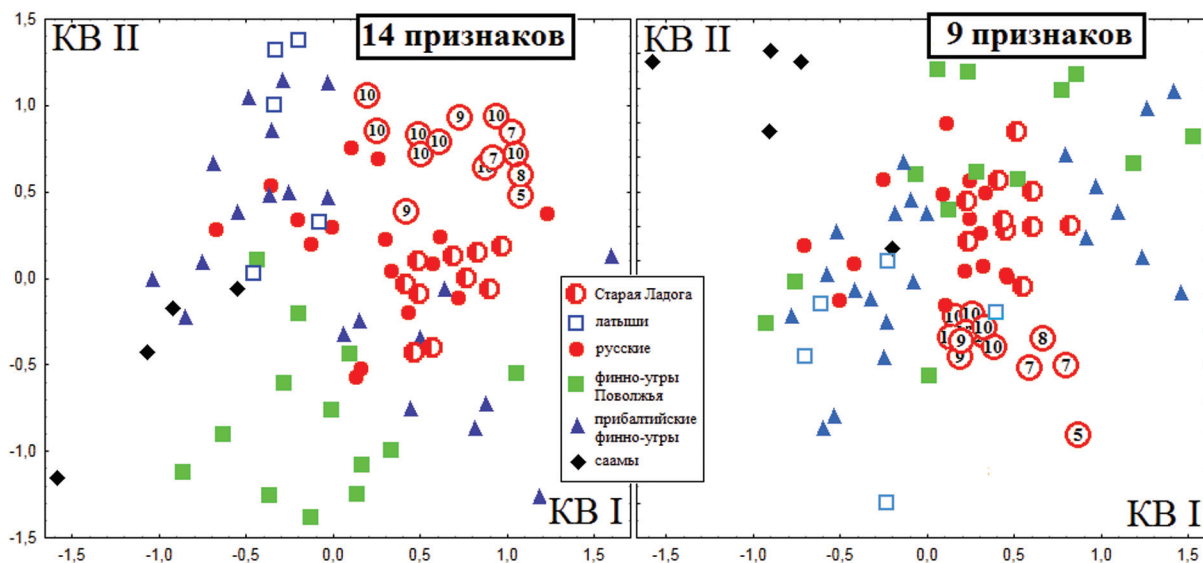


Рис. 4. Положение выборки русских Старой Ладого и некоторых краниологических серий Восточной Европы в пространстве первых двух канонических векторов

Примечания. Кружками с числами обозначены варианты положения выборки русских Старой Ладого по результатам измерений разных авторов; число в круге соответствует количеству черепов, вошедших в выборку.

при оценке соответствия стандартам полученных данных. Вероятно, подсказку для решения этого вопроса следует искать именно в оценке относительной согласованности изменения корреляций между компонентами и группами признаков, а не в сопоставлении ожидаемых и наблюдаемых единичных коэффициентов.

Как уже указывалось выше, разработка таких стандартов может стать предметом будущих исследований. При использовании существующих методов оценки однородности выборки, опирающейся на тестирование нормальности распределения первых главных компонент, в большинстве случаев исследователи, по-видимому, вполне могут доверять заключениям коллег.

Межгрупповой анализ

При оценке влияния межисследовательских расхождений на результаты межгруппового анализа учитывались измерения 14 признаков: 1, 8, 17, 9, 45, 48, 55, 54, 51, 52, 77, zm, SS:SC, 75(1). Группа этих признаков входит в стандартную программу, используемую в сравнительных исследованиях многими отечественными краниологами, и, в частности, используется в программе Б.А. Козинцева «CANON». Все выборки формировались с учетом различий в определении пола.

Кроме того, на основании индивидуальных измерений В.П. Алексева из общей выборки староладожских черепов случайным образом были сформированы 10 выборок численностью от 11 до 53 черепов. Последние использовались для визуальной сравнительной оценки эффектов, оказываемых межисследовательскими расхождениями и составом выборки, на результаты межгруппового сопоставления различных групп населения территории Восточной и Северо-Восточной Европы.

В качестве сравнительных материалов использовались данные по позднесредневековым и близким к современности сериям русских, латышей, финнов, эстонцев, саамов, карел, коми-пермяков, коми-зырян, удмуртов, марийцев и мордвы [Алексеев, 1969, 1971, 1974; Васильев, Боруцкая, 2004, 2011; Веселовская с соавт., 2015; Дубов, Дубова, 2000; Евтеев, 2011; Евтеев, Олейников, 2015; Комаров, Васильев, 2014; Пежемский, 2000, 2013; Харламова, 2012; Хартанович, 1986, 1991, 1995, 2004; Широбоков, 2014].

На рис. 4 слева приведены совмещенные результаты 24 вариантов канонического анализа, проведенного при помощи программы И.А. Гончарова «MultiCan». Единственное различие каждого из вариантов заключается в характеристике выборки из Старой Ладого: в 14 вариантах фигурирует одна выборка, измеренная разными исследователями, в 10 – случайные выборки из той же коллекции, измеренные В.П. Алексевым. По-

скольку в каждом варианте менялся состав лишь одной выборки, это позволило избежать заметных искажений в графическом отображении усредненного относительного расположения сравнительных групп.

Результаты, представленные на графике, могут быть сформулированы в виде четырех наблюдений:

1. В большинстве вариантов анализа наибольшее сходство выборки Старой Ладogi с остальными сериями русских является очевидным. Область значений, занимаемая староладожской выборкой в разных вариантах анализа, локализована в той же части графика, что и большая часть серий русских, хотя и занимает по отношению к ним периферийное положение в половине случаев.
2. Области, занимаемые разными случайными выборками, с одной стороны, и одной выборкой, но измеренной разными краниологами, не пересекаются.

Разное положение двух групп выборок связано не с отклонениями в измерениях В.П. Алексеева от других участников по ряду параметров, а с тем, что выборка, использованная для коннекции, не является случайной. Это обстоятельство открылось уже после завершения семинара. При помощи *t*-критерия Стьюдента было проведено тестирование различий между выборкой из 10 черепов и остальной частью коллекции (в обоих случаях сравнивались измерения В.П. Алексеева). Оказалось, что исследуемая выборка в целом массивнее, и на статистически значимом уровне отличается более широким лбом и сильно выступающим носом ($p < 0.05$).

При отборе материалов для измерений организаторы руководствовались одним принципом – черепа должны отличаться хорошей сохранностью, чтобы число измерений всех признаков, входящих в исследуемую программу, было максимальным. Может ли фактор лучшей сохранности более массивных черепов оказывать влияние на краниологическую характеристику серии в форме статистически значимых отклонений значений конкретных признаков? С полной уверенностью на этот вопрос ответить пока невозможно. Во всяком случае, мне не удалось выявить повторяющиеся значимые различия между размерами абсолютно целых и частично разрушенных черепов на примере 11 независимых выборок (саамов, русских, карел, финнов, шведов и коми-зырян). И все же, если обстоятельства вынуждают антрополога работать в поле с

массовым антропологическим материалом, предназначенным для перезахоронения, даже при ограниченном запасе времени, вероятно, лучшей тактикой будет не отбирать для измерения только черепа хорошей сохранности, а уделять внимание всем пригодным для краниологического анализа материалам с установленными половозрастными характеристиками.

3. Области значений, занимаемые разными случайными выборками, с одной стороны, и одной выборкой, измеренной разными краниологами, с другой, сопоставимы между собой по размаху.

Эффект, оказываемый на результаты межгруппового анализа межисследовательскими расхождениями в измерениях, сопоставим по своей величине с эффектом выборки. Очевидно, это не означает, будто одним из них можно пренебречь: влияние обоих эффектов накладывается друг на друга, а не нейтрализуется. Облако вероятных значений может быть выстроено вокруг координат каждой из сравниваемых выборок, а межисследовательские расхождения лишь увеличивают его размеры.

4. Область значений, занимаемая разными вариантами выборок Старой Ладogi, не уступает по размаху области, занимаемой всеми остальными сериями русских вместе взятыми.

Фактически такая высокая изменчивость результатов означает, что по результатам канонического анализа делать какие-либо выводы на основании сближения конкретных единичных выборок, совершенно некорректно. Разные исследователи, опираясь на один и тот же материал, с равной степенью обоснованности могут прийти к заключению, что: 1) характеристика староладожской серии по сравнению с большинством серий русских сдвинута в сторону отдельных прибалтийско-финских групп; 2) в староладожских черепках представлен типичный для русского населения краниологический комплекс; 3) при несомненном сходстве с другими сериями русских европейской части России русские Старой Ладogi отличаются рядом специфических черт. Приходится признать, что при изучении территориальной изменчивости населения в рамках отдельного географического региона надежные результаты, по всей вероятности, могут быть получены только при использовании данных собственных измерений.

К сожалению, перечисленные выше замечания остаются справедливыми, даже если провести более строгий подбор признаков (рис. 4, правый график). В новом анализе использовались те же 24 варианта староладожских серий, но кра-

ниометрическая программа была сокращена за счет исключения из нее признаков, при измерении которых возникают наибольшие расхождения. Использовались данные измерений 9 признаков: 1, 8, 17, 9, 45, 55, 54, 77, zm.

Казалось бы, область значений, занимаемая вариантами одной выборки, измеренной разными авторами, должна была значительно сузиться, однако этого не произошло. Уменьшение области разброса фиксируется лишь для выборок тех участников, определения половой принадлежности черепов у которых полностью или почти полностью совпадают между собой. При этом в относительном расположении выборок на графике проявляется отчетливая закономерность – чем больше между исследователями число различий в определении пола конкретных черепов, тем больше расстояния между соответствующими выборками.

Воспользовавшись табличными данными из работы В.П. Алексеева и Г.Ф. Дебеца [Алексеев, Дебеч, 1964], можно убедиться в том, что величина коэффициентов полового диморфизма (КПД) признаков, исключенных из анализа как ненадежных, превышает аналогичные параметры большей части оставшихся показателей. Однако в расхождениях результатов межгрупповых анализов важны не столько КПД используемых признаков, сколько соотношение КПД и коэффициента вариации. Влияние на сравнительную характеристику серии, оказываемое межисследовательскими расхождениями в определении пола, значительно сильнее проявляется в изменении размеров основных диаметров, нежели большинства признаков лицевого скелета, тем более при небольшой численности черепов. В первом анализе, проведенном для набора из 14 признаков, эффект пола выражен значительно слабее из-за наложения на него случайных отклонений в измерениях отдельных признаков, в частности, характеризующих форму орбит и степень выступления носовых костей.

Таким образом, изменение краниометрической программы за счет включения в нее только надежных (в отношении техники измерения) признаков, лишь отчасти снижает влияние межисследовательских расхождений на общие результаты сравнительно-популяционных исследований.

Заключение

По всей вероятности, в отношении оценки сопоставимости измерений конкретных признаков, полученные результаты в значительной степени соответствуют ожиданиям участников семинара.

В наибольшей степени варьируют признаки, характеризующие степень выступления носовых костей и переносья, а также вертикальные углы лицевого скелета, в меньшей степени – форма орбит и грушевидного отверстия. Наибольшей надежностью отличаются измерения основных диаметров, а также (как менее ожидаемый результат) углы горизонтальной профилировки.

При этом доверительные интервалы, в которых с 95% вероятностью находятся средние значения генеральных совокупностей, у отдельных исследователей совершенно не пересекаются между собой для четырех признаков: ширины орбиты, максиллофронтальной ширины, симотической ширины и среднего лицевого угла.

Вероятно, на серьезные расхождения данных признаков оказали влияние как случайные погрешности в измерении признаков малой величины, как в случае с симотической шириной (где, казалось бы, определение расстояния не должно вызывать разночтений), так и методические расхождения, связанные с трудностями локализации некоторых краниометрических точек, в частности, максиллофронтальной и назоспинальной. В других случаях, где также весьма вероятны методические расхождения, но измеряемые расстояния между точками больше в несколько раз, величина погрешности измерений, всё же, не превращает данные разных исследователей в абсолютно несопоставимые.

При этом общие результаты анализа главных компонент среднегрупповых значений, полученных участниками семинара, свидетельствуют, что фактор простой невнимательности при измерении и вводе данных вносит не меньший вклад в расхождения между исследователями, чем систематические методические расхождения. Человеческий фактор редко рассматривается как имеющий самостоятельное значение в анализе. Предполагается, что наиболее очевидные ошибки должны быть устранены еще до стадии статистической обработки, а мелкие случайные ошибки часто рассматриваются как менее опасные, чем систематические. К сожалению, реальная практика контроля случайных ошибок, по всей видимости, значительно отличается от ожидаемой. Об этом свидетельствуют явные ошибки в таблицах средних значений и показателей изменчивости признаков, периодически пропускаемые даже в рецензируемых журналах.

Кроме того, при работе с малыми выборками «положительное» отличие случайных ошибок от систематических никак себя не проявляет. Ожидаемой взаимной компенсации случайных отклонений чаще всего не происходит, а контролировать такие ошибки введением поправочных

коэффициентов и регрессионных формул невозможно.

В краткой форме влияние межисследовательских расхождений на результаты краниологических исследований может быть описано в форме следующих наблюдений:

- При условии корректного применения статистических методов результаты внутригруппового анализа (оценки однородности выборки), полученные разными исследователями, вероятно, могут считаться сопоставимыми, несмотря на расхождения в измерениях. По крайней мере, они сопоставимы в случаях, когда между исследователями отсутствуют значительные расхождения в определении половой принадлежности черепов.
- Суммарное влияние эффекта выборки и расхождений в измерениях на результаты межгрупповых сопоставлений весьма велико. Вероятно, заключения, основанные на сближении конкретных групп по результатам канонического анализа, не являются корректными не столько из-за искажения структуры близких связей в двумерных проекциях, сколько из-за суммарного вклада расхождений между исследователями и выборочного эффекта. Роль межисследовательских расхождений в снижении достоверности результатов, основанных на оценке минимальных расстояний (например, расстояний Махаланобиса), требует отдельного рассмотрения.
- Сокращение краниометрической программы за счет исключения из нее признаков, показавших наибольшую вариабельность между исследователями, не увеличивает существенно образом достоверность результатов межгруппового анализа из-за возрастающего влияния расхождений между исследователями в определении половой принадлежности черепов.
- В краниологическом исследовании, посвященном анализу населения конкретного географического региона, использование данных только собственных измерений и половозрастных определений имеет решающее значение при оценке достоверности полученных результатов;

Как бы ни были печальны некоторые итоги данного исследования, они никоим образом не должны использоваться в качестве аргумента для возвращения к визуальным оценкам морфологии черепа и применению типологического подхода к анализу краниологических материалов. Изменчивость морфологии черепа не может оцениваться иначе как при помощи корректно подобранных статистических инструментов. Безальтернатив-

ность такого подхода объясняется не только игнорируемым типологистами фактом значительного преобладания внутригрупповой изменчивости над межпопуляционной [Козинцев, 2016], но и тем обстоятельством, что только статистика позволяет контролировать ускользающую от верифицируемой оценки достоверность результатов исследователя, опирающегося на визуальные способности морфолога и опыт.⁴

Наиболее очевидный выход из сложившейся ситуации заключается во введении в качестве повседневной исследовательской практики повторного измерения всех признаков в каждой изучаемой серии, а также использовании сравнительных данных лишь тех исследователей, сопоставимость измерений которых не вызывает сомнений. Безусловно, задача снижения значения погрешностей, объясняющихся человеческим фактором, должна учитываться при разработке новых методов морфометрии, а их надежность в будущем может обеспечиваться свободным доступом к цифровым 3D-моделям. Контроль корректности интерпретации полученных результатов должен осуществляться при одновременном использовании данных разных систем признаков. В условиях возрастающего значения генетических исследований у краниологов еще сохраняется возможность продемонстрировать эффективность комплексного подхода в изучении истории формирования популяций прошлого.

Благодарности

Автор статьи выражает искреннюю признательность исследователям, принявшим участие в работе научно-практического семинара: А.Н. Багашеву, И.Р. Газимзянову, Н.Н. Гончаровой, А.В. Громову, А.А. Евтееву, А.В. Иванову, Д.С. Иконнико-

⁴Реальное значение последних умаляют сами типологисты. Несомненно, может существовать множество признаков, дифференцирующих популяции разного происхождения и не входящих в современную измерительную программу. Каждый такой признак может и должен быть формализован, а его изменчивость представлена хотя бы в форме дискретных вариантов. Именно при первичном отборе таких признаков решающую роль может сыграть глаз опытного специалиста-морфолога. Между тем, многие десятилетия отечественные краниологи используют в межгрупповых сопоставлениях одну и ту же стандартную программу, а редкие успешные попытки ее дополнения никак нельзя записать на счет сторонников типологического подхода.

ву, А.А. Казарницкому, Е.П. Китову, В.Г. Моисееву, Д.В. Пежемскому, В.И. Селезневой, К.Н. Солодовникову, Е.Н. Учаневой, В.И. Хартановичу и А.А. Хохлову.

Библиография

Алексеев В.П. Происхождение народов Восточной Европы (краниологическое исследование). М.: Наука, 1969.

Алексеев В.П. Очерк происхождения тюркских народов Восточной Европы в свете данных краниологии // Вопросы этногенеза тюркоязычных народов Среднего Поволжья. Казань, 1971. С. 232–271.

Алексеев В.П. Краниологическая характеристика населения Восточной Фенноскандии (по материалам Г.Ф. Дебеца и автора) // Расогенетические процессы в этнической истории. М., 1974.

Алексеев В.П., Дебеч Г.Ф. Краниометрия. Методика антропологических исследований. М., 1964.

Васильев С.В., Боруцкая С.Б. Палеоантропология населения Костромской области XVII в.: по материалам могильника Исупово // Расы и народы. 2004. Вып. 30. С. 249–267.

Васильев С.В., Боруцкая С.Б. Палеоантропологическое исследование первооселенцев г. Липецка (XVIII в.) // Актуальные вопросы антропологии. Минск: Беларуская навука, 2011. Вып. 6. С. 396–419.

Веселовская Е.В., Григорьева О.М., Пестряков А.П., Рассказова А.В. Антропологическая изменчивость населения Восточной и Центральной Европы от средневековья до современности // Вестник Московского университета. Серия XXIII. Антропология, 2015. №1. С. 4–24.

Дубов А.И., Дубова Н.А. Антропологическая характеристика четырех краниологических серий с территории Москвы // Народы России: от прошлого к настоящему. Антропология. Ч. 2. М.: Старый сад, 2000. С. 130–150.

Евтеев А.А. Краниологические материалы из некрополя с. Козино: предварительные результаты анализа // Актуальные вопросы антропологии. Минск: Беларуская навука, 2011. Вып. 6. С. 474–485.

Евтеев А.А., Олейников О.М. Археологические и палеоантропологические исследования на Даньславле улице в Великом Новгороде // Российская археология, 2015. № 1. С. 176–192.

Козинцев А.Г. О некоторых аспектах статистического анализа в краниометрии // Радловский сборник. Научные исследования и музейные проекты МАЭ РАН в 2015 году. СПб., 2016. С. 381–390.

Комаров С.Г., Васильев С.В. Краниологическое исследование группы лиц, погребенных в некрополях на территории Нижегородского Кремля // Вестник антропологии, 2014. Вып. 1 (27). С. 93–112.

Пежемский Д.В. Новые материалы по краниологии позднесредневековых новгородцев // Народы России. Антропология. Ч. 2. М.: Старый сад, 2000. С. 95–129.

Пежемский Д.В. Новые краниологические материалы по позднесредневековому населению Пскова // Вестник антропологии, 2013. Вып. 25. С. 121–126.

Пежемский Д.В., Харламова Н.В. Методический семинар по коннекции краниометрических программ // Вестник антропологии, 2013. Вып. 24. С. 169–172.

Харламова Н.В. Тверское население XVI–XX веков по данным краниологии // Вестник антропологии, 2012. Вып. 21. С. 49–58.

Хартанович В.И. Краниология карел // Антропология современного и древнего населения Европейской части СССР. Л.: Наука, 1986. С. 63–120.

Хартанович В.И. Новые материалы к краниологии коми-зырян // Сборник МАЭ, 1991. Т. 44. (Новые коллекции и исследования по антропологии и археологии). С. 108–125.

Хартанович В.И. Материалы к краниологии финнов // Антропология сегодня. Вып. 1. СПб., 1995. С. 71–89.

Хартанович В.И. Новые краниологические материалы по саамам Кольского полуострова // Палеоантропология, этническая история, этногенез. Сборник к 75-летию И.И. Гохмана. СПб.: МАЭ РАН, 2004. С. 108–125.

Широбоков И.Г. Материалы к антропологии пермских народов XVII–XIX вв. (краниологические данные) // Ежегодник финно-угорских исследований. Вып. 3. Ижевск: Изд-во «Удмуртский университет», 2014. С. 80–98.

Контактная информация:

Широбоков Иван Григорьевич: e-mail: ivansmith@bk.ru.

HOW SERIOUS IS AN EFFECT OF INTER-OBSERVER ERRORS ON CONCLUSIONS OF CRANIOLOGICAL STUDIES? (SOME RESULTS OF THE WORKSHOP ON CRANIOMETRIC CONSISTENCY IN PETER THE GREAT MUSEUM OF ANTHROPOLOGY AND ETHNOGRAPHY)

I.G. Shirobokov

Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), Russian Academy of Sciences, Saint Petersburg, Russia

In 2015 Peter the Great Museum of Anthropology and Ethnography (Kunstkamera) organized a workshop on estimation of inter-observer consistency in craniometry. Workshop participants measured two cranial series of Russians and Buryats. Each series consisted of 10 adult human skulls. The program of cranial measurements included traits widely used by anthropologists (№№ according to Martin: 1, 8, 17, 9, 11, 45, 48, 51, 52, 55, 54, SC, SS, MC, MS, 72, 73, 75(1)). The lowest consistency rates among participants were obtained for traits describing nasal bones and vertical facial angles, followed by the form of the orbit and piriform opening. The highest consistency rates were obtained for basic diameters and horizontal facial angles. Most probably, discrepancies in measurements are due to accidental errors in small size traits and methodical differences owing to difficulties in localizing some cranial landmarks (particularly maxillofrontale, ektokonchion and nasospinale). Mistakes made due to lack of attention are no less significant than methodological differences between anthropologists. The effect of inter-observer errors on results of intragroup analysis was evaluated using principal component analysis, on results of intergroup analysis – with discriminant canonical analysis. Results of intragroup analysis (homogeneity of sample) are comparable among researchers when there are no discrepancies in sex estimation. In intergroup analysis cumulative effect of inter-observer error and sample effect is very significant. It is not correct to evaluate similarity between given populations on the basis of rapprochement of corresponding samples within the first canonical vector space. Curtailment of craniometric program by means of removal of unreliable traits don't increase validity of research results because of inter-observer discrepancies in sex estimation. Solution to these problems may consist in introduction of compulsory practice of repeated measurement of all morphometric traits in each cranial series and in using comparative data of those colleagues whose method of measurement is known to be the same.

Keywords: craniometry, inter-observer consistency, intergroup analysis, intragroup analysis