



Ссылка на статью:

// Ученые записки УлГУ. Сер. Математика и информационные технологии. УлГУ. Электрон. журн. 2021, № 2, с. 1-13.

Поступила: 28.11.2021

Окончательный вариант: 01.12.2021

© УлГУ

УДК 57.017.6:[519.21+519.25]

О методологических и прикладных исследованиях в области биодемографии старения

Арбеев К. Г.

ka29@duke.edu

Университет Дьюка, Северная Каролина, США

В статье представлен обзор исследований и результатов автора и его коллег в Университете Дьюка (Северная Каролина, США) в области исследований процессов старения на основе стохастических моделей и анализа продольных данных. Описаны современные направления исследований по обнаружению генетических и негенетических детерминант старения, долголетия и здоровой продолжительности жизни и перспективы этой области исследований.

Ключевые слова: биодемография старения, стохастическая модель, анализ продольных данных, биомаркер

Введение.

От исследований финансовой математики к исследованиям старения

Моя научная карьера началась в области финансовой математики, когда я работал над кандидатской диссертацией на кафедре Прикладной Математики под руководством профессора А.А. Бутова в конце 1990-х гг. Диссертация была посвящена стохастическим моделям, предназначенным для исследования закономерностей в динамике активов на финансовых рынках (с конкретными приложениями к турбулентному российскому финансовому рынку середины и конца 1990-х годов). Это поистине увлекательная область исследований с множеством интересных и сложных теоретических разработок, основанных на стохастических процессах, а также важной практической реализацией этих теоретических моделей, которая привлекает и даёт возможность посвятить всю научную карьеру этой дисциплине. Однако мои научные исследования пошли другим путем, который может показаться совершенно не связанным с областью первоначальных исследований. Поворотным моментом стала моя работа в качестве приглашенного исследователя, а затем в качестве постдока в

Институте демографических исследований Макса Планка в Ростоке (Германия), который является одним из ведущих демографических центров и где мне посчастливилось сотрудничать с ведущими специалистами по демографии и другим смежным дисциплинам. Это дало мне уникальную возможность осознать, что исследования процесса старения – это тоже увлекательная область, в которой стохастические процессы могут успешно применяться при моделировании смертности, старения и развития болезней. С того времени мои основные научные интересы охватывают три широкие и взаимосвязанные дисциплины: биодемография, биостатистика и эпидемиология (в частности, генетическая эпидемиология). Я работаю в области исследований старения, которые сосредоточены на обнаружении генетических и негенетических детерминант старения, долголетия и здоровой продолжительности жизни, оцененных на основе так называемых «продольных» данных, и мои интересы связаны как с методологическими подходами в этой области исследований, так и с их практическим применением к анализу данных. Термин «продольные данные» (они же панельные данные, иногда их также называют лонгитюдными, от англ. longitudinal) означает данные, которые собираются посредством серии повторных наблюдений за одними и теми же субъектами в течение некоторого длительного периода времени. Результаты этой работы в настоящий момент представлены более, чем в 100 публикациях; часть из них составляет список литературы к данной статье, и они будут обсуждаться далее.

Результаты участия в исследовательских проектах в Институте демографических исследований Макса Планка легли в основу моей более поздней научной работы в Университете Дьюка в США. Первой публикацией в международном научном журнале стала статья, обобщающая результаты совместного проекта с участием профессора А.А. Бутова и еще одного выпускника УлГУ Д. Жданова [33]. Мы разработали математическую модель старения и смертности, которая позволила проверить гипотезу о том, что по крайней мере часть долгожителей могла иметь слабое здоровье в прошлом, и исследовали, как процессы индивидуального старения могут быть связаны с наблюдаемыми моделями показателей смертности. Кроме того, были проанализированы экспериментальные данные на лабораторных животных, чтобы выяснить, как соответствующие биомаркеры (например, масса тела) в более раннем периоде жизни могут быть связаны с более поздней смертностью и опухолевыми рисками у грызунов [1,2], а продольные данные о людях были использованы для изучения тенденций инвалидности в раннем пенсионном возрасте [8]. Это также была совместная работа с профессором А.А. Бутовым. Все эти проекты сформировали мои научные интересы, которые я продолжал расширять, работая в Университете Дьюка.

1. Биодемография старения:

научная дисциплина на пересечении биологии и демографии, требующая междисциплинарного мышления и навыков

Старение - сложный и многогранный процесс, развивающийся с возрастом. Следовательно, данные о различных соответствующих биомаркерах, измеренных в разном возрасте у одних и тех же людей, необходимы для получения знаний о механизмах и динамике

процесса старения. Потенциал и ценность биодемографических подходов в настоящее время осознаются благодаря доступности продольных измерений биомаркеров в существующих исследованиях и в тех, которые будут собирать продольные измерения биомаркеров в обозримом будущем. Возрастные траектории биомаркеров добавляют новое измерение к традиционному демографическому анализу, позволяя исследовать взаимосвязь динамики этих переменных и рисков смертности или заболеваемости и получить представление о возможных механизмах и динамике процессов, связанных со старением. Важная проблема, которую следует учитывать в контексте биодемографического анализа, заключается в том, что такой анализ направлен на объединение биологических знаний и теорий с демографическим анализом. В частности, в биодемографии старения это означает включение знаний и теорий о процессах старения в аналитические подходы. В научной литературе накоплен значительный объем знаний о механизмах изменений, связанных со старением, и сформулированы различные концепции старения. Концептуальная аналитическая структура, которая включает имеющиеся знания о механизмах изменений, связанных со старением, необходима для анализа их воздействия на риски смертности и заболеваемости. Разработка таких биодемографических моделей и их практическое применение для анализа данных является одним из основных направлений исследований нашего научного коллектива в Университете Дьюка.

1.1. Динамика биомаркеров в контексте старения, здоровья и смертности

Важно отметить, что научные открытия могут быть сделаны на основе эмпирических наблюдений и анализа собранных данных, а не только на основе сложного моделирования. Современные продольные исследования старения человека предоставляют множество возможностей для этого, и это было одним из направлений исследований нашей группы в последние десятилетия. В статье [22] мы выясняли, влияют ли формы возрастных траекторий некоторых физиологических переменных, таких как артериальное давление, индекс массы тела, глюкоза крови, холестерин, гематокрит и частота пульса в среднем возрасте (40–60 лет) на распределение остаточной продолжительности жизни. Мы продемонстрировали, что шансы на выживание в более старшем возрасте людей, доживших до 65 лет, действительно зависели от динамического поведения физиологических переменных в возрасте от 40 до 60 лет.

Люди, у которых значения биомаркеров существенно отклоняются от «оптимального» уровня, подвергаются более высокому риску смертности и, как правило, умирают раньше. Мы показали [27], что средние траектории различных биомаркеров у людей, умирающих в более раннем возрасте, существенно отклоняются от траекторий долгожителей: в первых группах траектории для многих биомаркеров имеют тенденцию увеличиваться до более высоких уровней и / или начинают снижаться раньше, иногда более быстрыми темпами, чем траектории тех кто живет дольше. Это указывает на то, что девиантная динамика биомаркеров с возрастом приводит к более высоким рискам смертности. Те же наблюдения можно сделать и для людей с короткой и длинной продолжительностью здоровой жизни [31]. Мы

обнаружили, что динамические характеристики (например, наклон, изменчивость) индивидуальных возрастных траекторий биомаркеров в среднем или пожилом возрасте влияют на выживаемость в более старшем возрасте [23, 27] и риски основных заболеваний, связанных со старением (рак, сердечно-сосудистые заболевания и диабет) [28].

Доступность генетической информации в продольных исследованиях позволяет рассчитывать средние возрастные траектории биомаркеров у носителей разных генотипов или аллелей. В работе [13] показано, что средние возрастные траектории холестерина и диастолического артериального давления у людей, умирающих в более раннем возрасте, заметно отклоняются от таковых в долгоживущих группах, и эти закономерности различаются для носителей и не носителей аллеля $e4$ гена $APOE$. Было замечено, что люди с большей продолжительностью жизни имеют стабильно более высокие уровни и менее резкое снижение как холестерина, так и диастолического артериального давления в пожилом возрасте (65+), когда такие уровни естественным образом снижаются в стареющем человеческом организме. Это наблюдение согласуется с выводами [23, 27] о том, что люди с самым медленным снижением биомаркеров имели лучшую выживаемость. Кроме того, было показано [29, 30, 34], что связанные со старением изменения физиологических переменных различаются у носителей и не носителей выбранных генетических вариантов, которые, как было установлено, связаны с продолжительностью жизни в полногеномных ассоциативных исследованиях. Удалось подтвердить, что генетическое влияние на продолжительность жизни реализуется через динамические механизмы, регулирующие изменения физиологических переменных в течение жизни, и что изменения физиологических переменных, вероятно, могут быть вызваны «скрытыми» (не наблюдаемыми непосредственно) компонентами старения, которые можно идентифицировать с помощью подходящей статистической методологии, которую я рассмотрю в следующем разделе.

Эти исследования демонстрируют важность изучения динамических характеристик биомаркеров в отношении старения, здоровья и смертности. Обобщение соответствующих результатов стало темой обзорной статьи [14]. Один из полученных в этой работе из выводов заключался в том, что, несмотря на обилие и растущее количество литературы по этой теме, потенциал анализа динамики биомаркеров в отношении смертности и заболеваемости с использованием строгих статистических подходов для совместного анализа времени до события и продольных наблюдений все еще остается в значительной степени недостаточно используемым, особенно в приложениях для кумулятивных мер, основанных на многих биомаркерах.

1.2. Методологические достижения в биодемографии старения: стохастическая модель процесса старения

Как отмечалось выше, объединение биологических знаний и теорий с демографическим анализом является краеугольным камнем дисциплины биодемографии старения. Формальные статистические подходы обычно бесполезны для этой цели, поскольку они не зависят от специфики исследовательской дисциплины и поэтому требуется разработка

конкретных биодемографических методов и моделей. Мы разработали подобную аналитическую структуру для этих целей – стохастическую модель процесса старения (далее СМПС) [24, 26], см. также введение в [14]). Отличительной особенностью этой модели является то, что она включает в себя доступные знания о механизмах изменений, связанных со старением, которые могут быть скрыты в индивидуальных продольных траекториях физиологических переменных, и это позволяет анализировать их косвенное влияние на риски заболеваний и смерти. В отсутствие конкретных наблюдений за продольной динамикой соответствующих биомаркеров, проявляющих такие механизмы и закономерности, традиционные статистические подходы имеют довольно ограниченную полезность для оценки соответствующих параметров, которые могут быть осмысленно интерпретированы с биологической точки зрения.

СМПС уходит своими корнями в модель случайного блуждания [21], которая примерно на десять лет предшествовала разработке моделей для совместного анализа времени до события и продольных наблюдений в биостатистической литературе (см. обзоры [4, 26]). С того времени СМПС была расширена различными способами и применена в различных контекстах для исследования механизмов связанных со старением в их связи с рисками заболеваемости и смертности. Например, модель была применена к данным о возрастных траекториях нескольких физиологических переменных и к данным о наступлении «нездоровой жизни» [11]. Было обнаружено, что динамические характеристики старения, такие как стрессоустойчивость, адаптивная способность, физиологическая «норма» (т.е. значения биомаркеров, минимизирующие соответствующий риск) и «аллостатические траектории» (т.е. значения, которым вынуждены следовать траектории биомаркеров под влиянием гомеостатических механизмов при наличии стрессов) изменяются с возрастом и различаются у женщин и мужчин. Важно отметить, что эти связанные со старением характеристики нельзя наблюдать непосредственно из данных, поскольку они скрыты в индивидуальных продольных траекториях физиологических переменных и могут быть проанализированы только косвенно с помощью статистического моделирования. Таким образом, СМПС является полезной моделью для оценки скрытых компонентов, связанных со старением, на основе продольных данных, которые не могут быть исследованы с использованием традиционных статистических методов. Хорошим источником информации об этой модели и других биодемографических методах и приложениях является книга [32], в которой обобщены недавние исследования нашей группы в области биодемографии старения, показывающие, как такие методы могут быть применены к имеющимся данным для поиска детерминант здоровой продолжительности жизни и долголетия.

Укажем три конкретных разработки в методологии СМПС, которые особенно актуальны в контексте последних достижений в исследованиях старения. Во-первых, это генетическая СМПС, специально предназначенная для совместного анализа генотипированных и негенотипированных участников продольных исследований [3-5, 15]. Генетическая СМПС позволяет оценивать (индивидуальные и плейотропные) генетические эффекты на такие не наблюдаемые напрямую характеристики, связанные со старением, как

физиологическая норма, аллостатическая адаптация, меры стрессоустойчивости и адаптационная способность и их влияние на риски смертности или заболеваемости. Такие оценки нельзя получить с помощью других известных подходов. Важно отметить, что это также дает дополнительную возможность для повышения точности генетического анализа путем совместного анализа данных для генотипированных и негенотипированных подвыборок по сравнению с традиционным анализом только генотипированных подвыборок [3, 15]. Во-вторых, версия СМПС со скрытой неоднородностью (латентными классами) [4, 5, 10, 25]. Этот подход помогает идентифицировать латентные субпопуляции с различными моделями продольных траекторий с различным влиянием на время до наступления события в каждой такой субпопуляции и различными паттернами скрытых связанных со старением характеристик, влияющих на эти признаки. Игнорирование такой скрытой структуры может привести к неверным оценкам и неверным интерпретациям; поэтому использование такой модели является полезным шагом в любом анализе, чтобы подтвердить надежность выводов из общих моделей. В-третьих, в [35] мы представили R-пакет *stm*, который является первым общедоступным набором утилит, реализующим методологию СМПС. Пакет позволяет оценивать несколько версий СМПС, доступных в настоящее время в литературе, включая многомерные модели с дискретным и непрерывным временем и одномерную модель с параметрами, зависящими от времени. Пакет доступен в репозитории CRAN и может использоваться бесплатно. Одним из основных прорывов, связанных с практическим использованием СМПС, была недавняя реализация в *stm* вычислительно эффективного алгоритма, который позволяет проводить крупномасштабный генетический анализ с использованием этой модели, как было проиллюстрировали в приложении к исследованиям ассоциаций в масштабе всего генома в [18].

Эти методологические достижения создали уникальную возможность использовать в значительной степени недостаточно используемый потенциал современных продольных данных, собирающих широкий спектр генетических данных, последующей информации о смертности и событиях, связанных со здоровьем, а также повторных измерений биомаркеров.

1.3. Составные биомаркеры в отношении старения, здоровья и смертности

Одно из важных направлений наших недавних исследований связано с объединением измерений нескольких биомаркеров в совокупные суммарные меры («составные биомаркеры») и их анализом в отношении процесса старения, здоровья и смертности. В настоящее время возможности применения современных статистических подходов для проведения совместного анализа динамики таких биомаркеров и времени до наступления события по-прежнему недостаточно используются [14]. Такие комплексные меры имеют преимущества как с технической точки зрения, так и с точки зрения предметной области. В вычислительном отношении их использование имеет явное преимущество, поскольку они позволяют работать в одномерном режиме вместо многомерного анализа исходных биомаркеров (хотя была разработана многомерная версия СМПС, см. обзор [26], её

рутинное использование в практических приложениях довольно ограничено, потому что вычислительные требования и технические трудности в многомерном случае часто делают проблемы трудноразрешимыми). С точки зрения применения в соответствующей области науки, их использование привлекательно, поскольку они позволяют анализировать системные и кумулятивные эффекты генетических и негенетических факторов на старение: они суммируют изменения в различных органах и физиологических системах в организме, которые могут накапливаться с течением времени сложным образом в различных регуляторных системах, но их кумулятивный эффект проявляется в увеличении физиологической дисрегуляции, которую можно количественно оценить с помощью таких кумулятивных мер, основанных на множестве биомаркеров. Кроме того, влияние изменений конкретных индивидуальных биомаркеров на соответствующие риски (смертность или заболеваемость) может быть небольшим, но совокупный эффект таких изменений в различных областях регуляторных систем может быть значительным, так что предсказывать их совместное влияние на процесс старения можно значительно лучше, чем влияние отдельных биомаркеров. В литературе появилось несколько методов построения кумулятивных (составных) биомаркеров, основанных на биологической теории, клинических данных и статистических соображениях (см., например, [16, 17, 19, 20] и ссылки в [14]). Эти методы позиционируются как способы количественной оценки биологического возраста индивида, который, в отличие от хронологического возраста (который является детерминированным фактором, изменяющимся с одинаковой скоростью для всех), может прогрессировать у различных индивидуумов с разной скоростью в результате воздействия различных генетических и негенетических факторов, накопленных в течение жизни.

Недавнее применение СМПС к составным биомаркерам [9] основывалось на статистическом расстоянии (дистанции Махаланобиса) (D_M), построенном на основе совместного распределения многих биомаркеров, которое можно интерпретировать как меру физиологической дисрегуляции [17]. Применение этой меры помогает избежать интенсивных вычислений, связанных с использованием многомерных моделей, и позволяет работать с несколькими коррелированными фенотипами в одномерных условиях. Мы обнаружили закономерности в физиологической дисрегуляции в связи с различными скрытыми механизмами изменений, связанных со старением, которые приводят к увеличению риска смерти с возрастом, а также могут быть связаны с наблюдаемыми гендерно-зависимыми различиями в рисках смертности.

В [6] было вычислено значение D_M с использованием измерений различных биомаркеров (связанными с воспалительными процессами, гематологическими, диабетическими, липидными, эндокринными, почечными) у участников Семейного Исследования Долгой Жизни (Long Life Family Study, LLFS). Обнаружено, что увеличение D_M было связано со значительно более высоким риском смертности, даже после поправки на составной показатель, суммирующий 85 нарушений, связанных со здоровьем (инвалидность, заболевания, менее серьезные симптомы), возраст и другие факторы риска. Оказалось, что такие

составные меры значительно улучшили прогнозы смертности, особенно в подвыборке участников из семей, характеризующимися исключительным долголетием. Эти результаты являются первым доказательством связи физиологической дисрегуляции со смертностью и ее прогностической эффективностью в уникальной выборке, отобранной на основании исключительного семейного долголетия.

Кроме того, оценивалось влияние D_M на общую выживаемость, начало «нездоровой» жизни (рак, сердечно-сосудистые заболевания или диабет; показатель «сопротивляемости организма») и выживаемость после начала «нездоровой» жизни (показатель «жизнестойкости организма») [12]. Выявлена более сильная связь физиологической дисрегуляции с наступлением «нездоровой» жизни, чем с выживанием после болезни. Это может означать, что снижение жизнестойкости и сопротивляемости с возрастом может иметь как пересекающиеся, так и различные механизмы. Мы пришли к выводу, что множественные отклонения физиологических маркеров от их нормального состояния (отраженные в более высоком значении физиологической дисрегуляции) могут способствовать повышению уязвимости ко многим заболеваниям и предшествовать их клиническому проявлению. Этот важный вывод подтверждает возможное использование этого составного показателя в здравоохранении в качестве доклинического индикатора перехода от здорового состояния к нездоровому. Эта работа также потребовала дальнейшего изучения генетических основ физиологической дисрегуляции и ее значимости для долголетия и возникновения болезней.

В [7] предложен новый подход к построению индивидуальных динамических характеристик меры физиологической дисрегуляции с использованием выходных данных совместных моделей выживания и продольных измерений биомаркеров, что, в частности, позволяет изучать ассоциации генетических вариантов с динамикой физиологической дисрегуляции. Мы выполнили общегеномное исследование ассоциации этих характеристик и обнаружили общегеномно-значимую ассоциацию генетического варианта на хромосоме 5 (расположенного в гене TRIO) с индивидуальными наклонами D_M . Применение разработанной методологии предоставило первые доказательства возможных генетических детерминант динамических характеристик физиологической дисрегуляции, которые основаны на статистически строгом подходе, позволяющем избежать систематической ошибки из-за незнания значений неслучайно отсутствующих измерений в традиционных методах (в данном случае это связано с неслучайным выбытием из-за смертности людей с более высоким уровнем физиологической дисрегуляции). Это важное методологическое развитие, потому что такие ситуации являются обычным явлением в исследованиях старения из-за самой природы изучаемых процессов. Изучение биологических эффектов генов, соответствующих наиболее значимым генетическим вариантам, обнаруженным в этом исследовании, показало, что эти гены широко участвуют в функции управления аксонами / синапсами. Это означает, что снижение способности аксонов поддерживать сложные регуляторные сети потенциально может играть важную роль в усилении физиологической дисрегуляции при старении. Полученные результаты закладывают основу для последующего анализа D_M в LLFS и других исследованиях.

2. О будущих направлениях биодемографии старения

Современная эпоха революционных достижений в генетике предоставляет большие возможности и вызовы для области биодемографии старения. Важность «генетической биодемографии» будет продолжать расти в ближайшие годы, потому что многие исследования, в которых были собраны данные о биомаркерах, будут включать (или уже включили) генетическую информацию. Подобное включение генетической информации в продольные исследования считается потенциально наиболее революционным элементом связанным с добавлением биологических данных в крупномасштабные исследования. Такие исследования будут во все большей степени предоставлять возможности для анализа взаимодействия генетических, биологических, социальных, экономических и демографических характеристик и их влияния на процесс старения, ухудшения здоровья и смертность.

Недавние крупномасштабные инициативы, такие как Британский биобанк (UK Biobank) и программа Trans-Omics for Precision Medicine (TOPMed), предоставляют беспрецедентные возможности для сбора фенотипической и генетической информации от сотен тысяч человек. В частности, подобные исследования предоставляют обширное генотипирование, полногеномное / экзомное секвенирование и другие данные по различным «-омикам» (протеомика, метаболомика и т.д.), которые могут помочь выяснить генетические и биологические основы различных заболеваний и нарушений здоровья. Дополнительной ценностью этих и других крупномасштабных исследований является то, что они собирают повторные измерения различных биологических и других переменных, а также последующие данные о смертности, возникновении заболеваний и других состояниях, связанных со здоровьем, для участников исследования. Такие исследования позволяют рассматривать перспективу всего жизненного цикла и оценивать индивидуальную возрастную динамику биомаркеров, которая отражает физиологическую дисрегуляцию в организме, и то, как такая динамика приводит к дифференцированным темпам ухудшения здоровья, различным рискам заболеваний и, в конечном итоге, различной продолжительности жизни. Наличие детальных генетических данных позволяет интегрировать генетические / «-омические» факторы в эту общую картину. Одна важная возможность, происходящая из масштаба таких данных – это возможность «индивидуализации» аналитических подходов. Такая «индивидуализация» может дать много преимуществ для выполнения более точных анализов и предсказаний в соответствии с современными тенденциями в отношении «персонализированной» медицины. Анализ таких богатых и разнообразных данных создает методологические проблемы из-за присущей им сложности и масштаба и требует соответствующих статистических инструментов. Быстро развивающаяся область машинного обучения и методов искусственного интеллекта должна привести к их более широкому применению в области исследований старения (включая биодемографию старения) для комбинированного анализа геномики и других многомерных данных, направленных на выявление детерминант долгой и здоровой жизни.

На сегодняшний день очевидна необходимость интеграции принципов генетики и геномики в биодемографию, и эта задача будет и впредь находиться в авангарде исследований

старения в ближайшие годы. Чтобы полностью реализовать потенциал богатого и быстро растущего объема данных, особое внимание следует уделять разработке методов анализа такой разноплановой информации. Мы ожидаем, что наши будущие и текущие исследования будут продолжать вносить вклад в эту развивающуюся программу исследований как в методологическом, так и в практическом плане.

Список литературы

1. Anisimov V.N., Arbeev K.G., Popovich I.G., Zabezhinski M.A., Arbeeve L.S., Yashin A.I. Is early life body weight a predictor of longevity and tumor risk in rats? // *Experimental Gerontology*. 2004, v. 39 (5), p. 807-816.
2. Anisimov V.N., Arbeev K.G., Popovich I.G., Zabezhinski M.A., Rosenfeld S.V., Piskunova T.S., Arbeeve L.S., Semenchenko A.V., Yashin A.I. Body weight is not always a good predictor of longevity in mice // *Experimental Gerontology*. 2004, v. 39 (3), p. 305-319.
3. Arbeev K.G., Akushevich I., Kulminski A.M., Arbeeve L.S., Akushevich L., Ukraintseva S.V., Culminskaya I.V., Yashin A.I. Genetic model for longitudinal studies of aging, health, and longevity and its potential application to incomplete data // *Journal of Theoretical Biology*. 2009, v. 258 (1), p. 103-111.
4. Arbeev K.G., Akushevich I., Kulminski A.M., Ukraintseva S., Yashin A.I. Joint analyses of longitudinal and time-to-event data in research on aging: Implications for predicting health and survival // *Frontiers in Public Health* 2, 2014, article 228.
5. Arbeev K.G., Arbeeve L.S., Akushevich I., Kulminski A.M., Ukraintseva S.V., Yashin A.I. Latent Class and Genetic Stochastic Process Models: Implications for Analyses of Longitudinal Data on Aging, Health, and Longevity // In: "*JSM Proceedings, Section on Statistics in Epidemiology*." American Statistical Association: Alexandria, 2015. VA, p. 121-133.
6. Arbeev K.G., Bagley O., Ukraintseva S.V., Duan H., Kulminski A.M., Stallard E., Wu D., Christensen K., Feitosa M.F., Thyagarajan B., Zmuda J.M., Yashin A.I. Composite Measure of Physiological Dysregulation as a Predictor of Mortality: The Long Life Family Study // *Frontiers in Public Health*. 2020, v. 8, article 56.
7. Arbeev K.G., Bagley O., Ukraintseva S.V., Wu D., Duan H., Kulminski A.M., Stallard E., Christensen K., Lee J.H., Thyagarajan B., Zmuda J.M., Yashin A.I. Genetics of physiological dysregulation: findings from the long life family study using joint models. // *Aging*. 2020, v. 12 (7), p. 5920-5947.
8. Arbeev K.G., Butov A.A., Manton K.G., Sannikov I.A., Yashin A.I. Disability trends in gender and race groups of early retirement ages in the USA // *Sozial Und Praventivmedizin*/ 2004, v. 49 (2), p. 142-151.
9. Arbeev K.G., Cohen A.A., Arbeeve L.S., Milot E., Stallard E., Kulminski A.M., Akushevich I., Ukraintseva S., Christensen K., Yashin A.I. Optimal versus Realized Trajectories of Physiological Dysregulation in Aging and their Relation to Sex-Specific Mortality Risk // *Frontiers in Public Health*. 2016, v. 4, article 3.

10. Arbeev K.G., Land K.C., Yashin A.I. The Latent Class Stochastic Process Model for Evaluation of Hidden Heterogeneity in Longitudinal Data // In: "*Biodemography of Aging: Determinants of Healthy Life Span and Longevity.*" Springer Netherlands: Dordrecht, 2016. P. 285-302.
11. Arbeev K.G., Ukraintseva S.V., Akushevich I., Kulminski A.M., Arbeeveva L.S., Akushevich L., Culminskaya I.V., Yashin A.I. Age trajectories of physiological indices in relation to healthy life course // *Mechanisms of Ageing and Development.* 2011. v. 132 (3), p. 93-102.
12. Arbeev K.G., Ukraintseva S.V., Bagley O., Zhbannikov I.Y., Cohen A.A., Kulminski A.M., Yashin A.I. "Physiological Dysregulation" as a Promising Measure of Robustness and Resilience in Studies of Aging and a New Indicator of Preclinical Disease // *Journals of Gerontology Series A Biological Sciences and Medical Sciences.* 2019, v. 74 (4), p. 462-468.
13. Arbeev K.G., Ukraintseva S.V., Kulminski A.M., Akushevich I., Arbeeveva L.S., Culminskaya I.V., Wu D., Yashin A.I. Effect of the APOE Polymorphism and Age Trajectories of Physiological Variables on Mortality: Application of Genetic Stochastic Process Model of Aging // *Scientifica.* 2012, Article ID 568628.
14. Arbeev K.G., Ukraintseva S.V., Yashin A.I. Dynamics of biomarkers in relation to aging and mortality // *Mechanisms of Ageing and Development.* 2016, v. 156, p. 42-54.
15. Arbeev K.G., Yashin A.I. How Biodemographic Approaches Can Improve Statistical Power in Genetic Analyses of Longitudinal Data on Aging, Health, and Longevity // In: "*Biodemography of Aging: Determinants of Healthy Life Span and Longevity.*" Springer Netherlands: Dordrecht, 2016. P. 303-319.
16. Belsky D.W., Caspi A., Houts R., Cohen H.J., Corcoran D.L., Danese A., Harrington H., Israel S., Levine M.E., Schaefer J.D., Sugden K., Williams B., Yashin A.I., Poulton R., Moffitt T.E. Quantification of biological aging in young adults // *Proceedings of the National Academy of Sciences of the United States of America.* 2015, v. 112 (30), E4104-E4110.
17. Cohen A.A., Milot E., Yong J., Seplaki C.L., Fuloep T., Bandeen-Roche K., Fried L.P. A novel statistical approach shows evidence for multi-system physiological dysregulation during aging // *Mechanisms of Ageing and Development.* 2013, v. 134 (3-4), p. 110-117.
18. He L., Zhbannikov I., Arbeev K.G., Yashin A.I., Kulminski A.M. A genetic stochastic process model for genome-wide joint analysis of biomarker dynamics and disease susceptibility with longitudinal data // *Genetic Epidemiology.* 2017, v. 41 (7), p. 620-635.
19. Howlett S.E., Rockwood M.R.H., Mitnitski A., Rockwood K. Standard laboratory tests to identify older adults at increased risk of death // *BMC Medicine.* 2014, 12: 8.
20. Seeman T.E., Singer B.H., Rowe J.W., Horwitz R.I., McEwen B.S. Price of adaptation - Allostatic load and its health consequences: MacArthur studies of successful aging // *Archives of Internal Medicine.* 1997, v. 157 (19), p. 2259-2268.
21. Woodbury M.A., Manton K.G. A random-walk model of human mortality and aging // *Theoretical Population Biology.* 1977, v. 11 (1), p. 37-48.

22. Yashin A.I., Akushevich I.V., Arbeev K.G., Akushevich L., Ukraintseva S.V., Kulminski A. Insights on aging and exceptional longevity from longitudinal data: novel findings from the Framingham Heart Study // *Age*. 2006, v. 28 (4), p. 363-374.
23. Yashin A.I., Arbeev K.G., Akushevich I., Arbeeveva L., Kravchenko J., Il'yasova D., Kulminski A., Akushevich L., Culminskaya I., Wu D., Ukraintseva S.V. Dynamic determinants of longevity and exceptional health // *Current Gerontology and Geriatrics Research*. 2010, 381637.
24. Yashin A.I., Arbeev K.G., Akushevich I., Kulminski A., Akushevich L., Ukraintseva S.V. Stochastic model for analysis of longitudinal data on aging and mortality // *Mathematical Biosciences*. 2007, v. 208 (2), p. 538-551.
25. Yashin A.I., Arbeev K.G., Akushevich I., Kulminski A., Akushevich L., Ukraintseva S.V. Model of hidden heterogeneity in longitudinal data // *Theoretical Population Biology*. 2008, v. 73 (1), p. 1-10.
26. Yashin A.I., Arbeev K.G., Akushevich I., Kulminski A., Ukraintseva S.V., Stallard E., Land K.C. The quadratic hazard model for analyzing longitudinal data on aging, health, and the life span // *Physics of Life Reviews*. 2012, v. 9 (2), p. 177-188.
27. Yashin A.I., Arbeev K.G., Ukraintseva S.V., Akushevich I., Kulminski A. Patterns of Aging Related Changes on the Way to 100: An Approach to Studying Aging, Mortality, and Longevity from Longitudinal Data // *North American Actuarial Journal*. 2012, v. 16 (4), p. 403-433.
28. Yashin A.I., Arbeev K.G., Ukraintseva S.V., Arbeeveva L.S., Akushevich I., Kravchenko J., Kulminski A.M., Culminskaya I., Wu D., Land K.C. Dynamic Characteristics of Aging-Related Changes as Predictors of Longevity and Healthy Lifespan // In: "*Biodemography of Aging: Determinants of Healthy Life Span and Longevity*." Springer Netherlands: Dordrecht, 2016. P. 187-210.
29. Yashin A.I., Arbeev K.G., Wu D., Arbeeveva L., Kulminski A., Kulminskaya I., Akushevich I., Ukraintseva S.V. How Genes Modulate Patterns of Aging-Related Changes on the Way to 100: Biodemographic Models and Methods in Genetic Analyses of Longitudinal Data // *North American Actuarial Journal*. 2016, v. 20 (3), p. 201-232.
30. Yashin A.I., Arbeev K.G., Wu D., Arbeeveva L.S., Kulminski A., Akushevich I., Culminskaya I., Stallard E., Ukraintseva S. How lifespan associated genes modulate aging changes: lessons from analysis of longitudinal data // *Frontiers in Genetics*. 2013, v. 4, article 3.
31. Yashin A.I., Arbeeveva L.S., Arbeev K.G., Akushevich I., Kulminski A.M., Stallard E., Ukraintseva S.V. Age Trajectories of Physiological Indices: Which Factors Influence Them? // In: "*Biodemography of Aging: Determinants of Healthy Life Span and Longevity*." Springer Netherlands: Dordrecht, 2016. P. 21-45.
32. Yashin A.I., Stallard E., Land K.C. *Biodemography of Aging: Determinants of Healthy Life Span and Longevity*. Springer: New York, 2016.
33. Yashin A.I., Ukraintseva S.V., De Benedictis G., Anisimov V.N., Butov A.A., Arbeev K., Jdanov D.A., Boiko S.I., Begun A.S., Bonafe M., Franceschi C. Have the oldest old adults

ever been frail in the past? A hypothesis that explains modern trends in survival // *Journals of Gerontology Series A Biological Sciences and Medical Sciences*. 2001, v. 56 (10), B432-B442.

34. Yashin A.I., Wu D., Arbeeveva L.S., Arbeev K.G., Kulminski A.M., Akushevich I., Kovtun M., Culminkaya I., Stallard E., Li M., Ukraintseva S. Genetics of aging, health and survival: Dynamic regulation of human longevity related traits // *Frontiers in Genetics*. 2015, v. 6, article 122.
35. Zhbannikov I.Y., Arbeev K., Akushevich I., Stallard E., Yashin A.I. STPM: an R package for stochastic process model // *BMC Bioinformatics*. 2017, v. 18 (1): 125.

Methodological and applied research in biodemography of aging

Arbeev, K. G.

ka29@duke.edu

Duke University, North Carolina, USA

The paper provides an overview of research and results of the author and his colleagues at Duke University (North Carolina, USA) in the field of research on aging based on stochastic models and analysis of longitudinal data. The author also describes current directions of research on evaluation of genetic and non-genetic determinants of aging, longevity, and healthy life span, as well as future prospects in this research area, are described.

Keywords: *biodemography of aging, stochastic model, analysis of longitudinal data, biomarker.*