

МОДУЛЬ 2. Управление данными и методология

2.2 Развитие новых технологий в спорте, новые данные

Marti Casals

В последние годы произошли важные изменения в спорте, что не только создало большие ожидания, но и оказало влияние на экономику, здоровье и подписания игроков. Более того, это генерирует новую область знаний. Мы перешли от записей в блокноте при просмотре футбольного или баскетбольного матча (анализ рукописных данных) к распоряжению огромным количеством данных, собираемых высокотехнологичной аппаратурой.

Ныне популярные наручные браслеты, которые часто носят на себе спортсмены во время тренировок и матчей имеют сенсоры для сбора медицинской информации в области суставов, где они закрепляются. В последние годы было проделано множество работы в области координации и кросс-совместимости информации из разного программного обеспечения. Сейчас тактические, физические и медицинские аспекты интегрированы и централизованы в облачные сервисы. В ближайшем будущем это позволит достигнуть мультидисциплинарности при работе множества специалистов (менеджеры, тренеры, доктора, физические терапевты, тренеры по ФП, аналитики), которые будут интегрированы в работе спортивной индустрии.

В наши дни на спорт оказывает влияние большое количество анализируемой информации, широкое разнообразие окружающей среды от физиологии упражнений, психологии и продукции тепла до современных систем анализа спортивных показателей в течение соревнований. Разные профессиональные команды начинают собирать информацию по контролю за травмами и осмотром спортсменов с помощью разных устройств. Но как эти онлайн-системы контроля за здоровьем спортсменов воспринимаются ими самими или персоналом команд? Принимаются или отрицаются? В соответствии с исследованием (Barboza, Bolling, Nauta, Van Mechelen, and Verhagen, 2017) они помогают улучшить коммуникацию между тренерским штабом, спортсменом и медицинским персоналом. Более того, с точки зрения врача они помогают более глубоко понимать и контролировать спортсмена и вмешиваться когда это необходимо (Barboza, et al., 2017).

Одной из главных проблем при использовании такого количества технологичных устройств будет определить какой из приборов может выдавать ошибочные результаты. В спорте многие диагностические мероприятия проводятся при использовании плохо воспроизводимых методов, хотя и широко используются. Можем ли мы быть уверены, что параметры, записанные во время обследования достоверны? И зависит ли эта достоверность - или ее недостаток - от того, кто проводит тесты? Нет сомнений, что мы принимаем решения исходя из этих тестов, но являются ли данные от этих тестов корректными? Наша ли это ответственность? Это вопросы, которые мы должны задать себе сами.

В наши дни количество получаемых данных постоянно растёт и продолжит расти еще больше вместе с технологическим прогрессом. Такая доступность данных заставляет очень щепетильно подбирать самые достоверные, создавать индикаторы для дальнейшего приобретения знаний, и, надо всем этим, создавать статистические модели, касающиеся производительности спортсменов, их адаптации и риска травм. Для членов спортивного сообщества, близких к научной сфере, абсолютно понятно что анализ и статистическая и математическая обработка огромного количества получаемых данных открывает новые области знаний и простор для бизнеса. Подводя итог можно сказать, что у нас сейчас очень много технологий в нашем распоряжении, но нам всё ещё нужно работать над принятием более правильных решений, основанных на адекватных моделях и статистических анализах.

2.2.1 Статистика и спорт

Ходит много разговоров о статистике в спорте. Это сопровождается бумом статистиков как профессии и как науки. Быть аналитиком - одна из самых сексуальных профессий 21го века. На самом ли деле это стоящая тема или всё-таки прихоть? История показала, что это не так. Важно отметить Henry Chadwich с его нотационным анализом еще из 1861; профессор James Connors составил первую тепловую карту в спорте в 1897; Mr. Roth был первым нанятым на полное рабочее время статистиком в профессиональной команде - это произошло в 1947. Список можно продолжить, упомянув, например, известную книгу Michael Lewis под названием "Moneyball: the art of winning an unfair game" (Манибол: искусство побеждать в нечестной игре), опубликованную в 2003. Книга повествует реальной истории Билли Бина, генерального менеджера скромной Калифорнийской команды, решившего использовать статистический подход для принятия решений - его роль сыграл Брэд Питт.

Очень легко представить кто будет заинтересован в спортивной статистике. Число профессионалов, заинтересованных этой областью поражает: спортивные менеджеры, люди принимающие решения в клубах, игроки, тренеры, тренеры ФП, врачи,

журналисты, писатели, скауты, аналитики, учёные, академики, фанаты, психологи, эпидемиологи, различные эксперты и др.

Существенно возрос интерес к приложению статистики к науке и её возросшая точность в области спортивной медицины и спортивной науки в целом. Статистики, вместе со специалистами в других областях, установили основы для разных специализаций, таких как биостатистика, биоинформатика, геостатистика, эконометрика и психометрика (Схема 1). Голливудский фильм “Человек, который изменил всё” также помог поднять интерес к науке в спорте, в особенности к саберметрике, спортивной аналитике и спортивной биостатистике, которые берут своё начало от статистики (Схема 1).

Таблица 1: Общие специализации в области статистики

Наиболее частые специальности в области статистики
Статистика и Эпидемиология, и Здоровье Нации Медицина ≈ Биостатистика Статистика и Компьютерная Наука, Биология и Генетика ≈ Биоинформатика Статистика и География ≈ Геостатистика Статистика и Психология ≈ Психометрика Статистика и Экономика ≈ Эконометрика
Специализации по статистике и спортивной медицине
Статистика и Бейсбол ≈ Саберметрика (Sabermetrics) Статистика и Спортивная наука, Экономика и Компьютерная Наука ≈ Moneyball Статистика и Спортивная Наука, Видео-аналитика и Компьютерная Наука ≈ Спортивный аналитик Статистика и Эпидемиология, и Здоровье Нации Медицина и Компьютерная Наука ≈ Спортивный Биостатистик

Источник: Casals y Finch, 2017

2.2.2 Что такое статистика и насколько она полезна?

Многие из нас проходили курс по Статистике в университете, но никто особо не помнит что это такое и почему это так полезно. Статистика более не является просто ответвлением Математики, а стала наукой об обучении на основе собранных данных, которая измеряет и контролирует неопределенность.

Как наука она очень молода и многие люди еще не знают о её существовании. Доказательством этого является, что статистиков принимают за государственных служащих или думают, что о них, как о каких-то ученых, выполняющих вычисления или считающих средние значения.

Статистика количественно определяет неопределенность и создает рекомендации как собирать данные, чтобы они давали как можно больше информации. Биостатистика - наука о статистике в науках о жизни, в особенности о здоровье - изучается в медицине. Большая разница между статистикой и биостатистикой состоит в том, что нужно разбираться еще и в других направлениях типа эпидемиологии, генетике, демографии, здоровье населения и тд.

Когда мы читаем научные статьи или хотим заняться научным проектом важно понимать некоторые основные правила:

- как различать популяцию (цель нашего исследования и все данные, которые мы хотим собрать) от образца (группа или часть популяции и данные, с которыми мы можем работать).

Также важно знать как выбирать образцы и разные техники отбора (например рандомизированную, стратифицированную), причины их выбора или информационные отклонения.

- чтобы проводить точный анализ обязательно знать как дифференцировать природу, уровень измерений, шкалу и тип вариабельности (качественную или количественную) или характеристики, на которых мы хотим построить наше исследование.
- Правильно пользоваться статистикой значит знать на какие вопросы мы хотим получить ответ. Вопросы типа “что происходит?”(описательная статистика), “что случилось?” (диагностическая статистика), “что случится?” (предсказательная статистика) или “что нам следует делать?” (директивная статистика) указывают нам на разные статистические инструменты. Другим ключевым отличие, которое надо держать в уме - это то, что в исследованиях могут изучаться полученные данные (описательная статистика) или делать выводы о популяции (дедуктивная статистика). Описательная статистика показывает нам данные, которые есть у нас в определенный момент с помощью, например, таблиц частот и графиков в соответствии с переменными, которые мы изучаем. Для измеряемых величин используются разные измерения, например центральность, дисперсия, позиция и форма. Например в Университете Северной Каролины в Chapel Hill привыкли хвастаться, что их студенты Географии получают в среднем более высокую зарплату после окончания в отличие от студентов других университетов. Чего они не знали, так это того, что среди студентов того направления были те, кто учится по спортивной стипендии. Этим студентом был Майкл Джордан, чей доход был значительно выше остальных выпускников. В этом случае мы должны считать не среднюю зарплату, а использовать метод центральности, такой как вычисление

медианы. Alberto Cairo, инфограф, пишет: “статистика не врет, врет тот, кто ей манипулирует”. Это также будет важным моментом при соединении измерения центральности с каким-нибудь измерением дисперсии для выявления надежности данных. При описании измеряемых переменных полезно использовать box-plot - это стандартизированный способ отображения набора данных на основе пяти показателей: минимума, максимума, медианы выборки и первого и третьего квартилей, и который позволяет нам раскрыть резко отличающиеся значения. Давайте представим, что биостатистик сообщает главному врачу клуба, что нынешний 3й перцентиль травм плеча составляет 7 случаев. Как эксперты в спортивной медицине, мы можем интерпретировать, что 3% спортсменов в нашем клубе будут иметь 7 или менее травм плеча. Обладание навыками коммуникации и интерпретации этих базовых измерений обязательно и необходимо.

- Статистические выводы направлены на то, чтобы делать заключение о популяции на основе анализа группы. Должны быть освещены два концепта: статистически важные отличия и клинически релевантные отличия. Какой из них более важен? Возможно второй, но чтобы ответить на этот вопрос нам нужен и первый. Мы живем в мире, где постоянно сравниваем себя. Отправная точка для сравнения, соотношения, тестирования или оценки - это проверка гипотез. Также существует концепция моделирования. Основой многих научных публикаций является концепция клинической значимости, обычно оцениваемая индексом, называемым значением P (пи). Значение P - это вероятность того, что основанная на определенной статистической модели, статистически суммирующая некоторые характеристики данных (например отличия в значениях при сравнении двух групп)...

Однако, несмотря на то, что значение P - очень полезный статистический показатель, он часто неправильно используется и интерпретируется. Это разубедило использовать его многими научными журналами. В этом свете Американская Статистическая Ассоциация (ASA) сделала официальное заявление для научного сообщества, разъяснив некоторые общепринятые принципы, которые подразумеваются при правильном использовании и интерпретации значения P (Американская статистическая ассоциация, 2016). Одиночный индекс не должен подменяться в научных рассуждениях. Также были предложены некоторые альтернативы этому индексу (такие как доверительный интервал [CI], размер эффекта, D Коэна [Cohen's D] и тд.) и их нужно держать в уме.

- Когда мы соотносим две числовые переменные, которые имеют линейную зависимость, говорят, что эти переменные имеют корреляцию. Эта концепция (прямой зависимости) часто путается с причинно-следственной связью, немного другой концепцией, имеющей другой дизайн и для которой необходимо учитывать другие аспекты.
- Часто целевую или зависимую переменную сравнивают с другой интересующей характеристикой, или независимой переменной. Однако, в спортивной медицине, как и в других сферах, главный объект (например: травмироваться - да/нет) часто ассоциируется более чем с одной ковариацией (например: предыдущие травмы, возраст, игровое покрытие, время сезона и тд) и множеством факторов - некоторые скрытые, некоторые динамические - те, которые могут влиять на это, должны быть обязательно рассмотрены. Спецификация и утверждение статистических регрессивных моделей, и включение сбивающих с толку переменных - это аспекты, которые должны быть проверены для понимания этиологии травм или факторов с ними ассоциированных, как и использование предсказательных моделей.
- Большинство вопросов спортивной медицины, на которые хотят получить ответ специалисты - диагноз, прогноз и поиск факторов риска травм или болезней. Это могут быть изменяемые факторы (например если атлет курит это может повышать риск определенных заболеваний) и неизменяемые (например возраст). В спортивной медицине также часто говорят о внутренних факторах (возраст, этническая принадлежность, психологические факторы, баланс мышечной силы, тренировочная нагрузка, гибкость, утомление и тд). Врачи часто пытаются учитывать все из них, и нет ничего лучше статистической модели, которая помогла бы им в этом. Последнее десятилетие ознаменовано проделанной биостатистиками работой по персонализированному подходу в медицине, которые помогает назначить пациенту правильную дозу лечения, основываясь на их уникальных характеристиках, при этом лечение одинаковых патологий может отличаться (goo.gl/gwCUGD). Однако, знание факторов риска или выявление подгрупп, имеющих более высокий риск травм, не всегда достаточно. Важно выяснить откуда травмы и заболевания берутся.

2.2.3 Методология исследований и эпидемиология тренировок

Профессионалы в области наук о жизни не могут недооценивать значимость исследований, и спортивная медицина не является исключением. Как в любой науке знания добываются с помощью научного метода. Чтобы провести исследование важно иметь возможность задавать вопросы и следовать научному методу. Интуитивные

знания или неподтвержденные догадки просто не подходят. Научный метод - это последовательность этапов (наблюдение, введение, гипотеза; индукция, доказательство или проверка гипотезы экспериментом; демонстрация или опровержение гипотезы; и тезис или научная теория), которому ученые следуют для того, чтобы делать новые открытия и проверять до сих пор неизвестные гипотезы. Исследование и этапы научного метода в основном соотносятся с часто отображаемыми частями научных статей (смотри таблицу 1) (Mabrouki, and Bosch, 2007).

Таблица 2: Этапы научного метода

Этапы научного метода	Основные разделы научной статьи
Понимание изучаемой проблемы	Введение
Выдвижение гипотезы	
Сбор данных	Материалы и методы
Анализ данных	Результаты
Интерпретация результатов	Обсуждение
Заключение / выводы	

Источник: Собственная разработка автора

Ключевым аспектом при проведении исследования будет знание как задавать правильные вопросы, следовать научному методу и знать разные эпидемиологические дизайны, показатели (частота, ассоциация, влияние), и следовать одобренным научным гайдлайнам, таким как STROBE (улучшение отчетности наблюдательных исследований в эпидемиологии), CONSORT (объединенные стандарты отчетности об испытаниях), PRISMA (предпочтительные элементы отчетности для системных обзоров и мета-анализа), и WHO Injury Surveillance Guidelines (рекомендации по наблюдению за травмами Всемирной Организации Здравоохранения). Эпидемиология - наука, изучающая распределение и причины болезней и травм в определенной популяции. В индустрии спорта Международный Олимпийский Комитет работает, чтобы укрепить здоровье спортсменов на годы вперед с помощью специальных научных эпидемиологических групп по всему миру.

В спортивной медицине существуют различные профили, которые включают разные навыки и цели. Все профили, описанные ниже нужны клубам, в дополнение к интересу, энтузиазму и страсти к спортивной медицине.

- клинический профиль: прием и диагностика пациентов каждый день не должны отдалять врача от клинической практики и реальности спортивной медицины.
- эпидемиологический профиль: сбор информации о травмах, изучение их распределения и возможных причин помогает предотвращать их и проводить мониторинг спортсменов.
- научный профиль: написание и понимание научных статей, следуя клиническим практическим рекомендациям, делает специалиста ближе к преимуществам спортивной медицины.
- профиль видеоанализа: просмотр матчей и тренировок онлайн или на ТВ и попытка понять эволюцию спорта помогает лучше ставить клинические или научные задачи.
- профиль статистика/аналитика: попытка понять паттерны и непредсказуемость событий поможет иметь более удобные для подсчета результаты.

Споры вокруг того, что кто-то является в большей степени клиницистом или эпидемиологом бесполезны, так как оба навыка нужны в современной медицине. Сейчас не всегда существует разделение между клинической практикой и исследованиями, между клиницистами и учеными. Даже при том, что не все врачи занимаются научной работой, мы должны быть активными пользователями научной литературы, которая предоставляет нам множество знаний и добавляет точность в действиях. В последнее время это разделение становится всё менее заметным благодаря большому количеству мультидисциплинарной работы и потребности знать как интерпретировать большое количество накопленных данных. Британский Журнал Спортивной Медицины (BJSM) в данный момент занимается написанием нескольких образовательных статей, освещающих статистику и эпидемиологию. Это делает работающих в командах людей ближе к научному сообществу (Nielsen, et al., 2017a; Nielsen, et al., 2017b). Эти статьи помогают тренерам и врачам знать как читать научные статьи и различать понятия, такие как, например, распространённость и частота травм. Разные ученые пытаются улучшить различные аспекты как в других медицинских дисциплинах. John Ioannidis (2005), один из пионеров так называемой мета-науки, дисциплины, анализирующей работу других ученых, определяя по фундаментальным правилам какая из научных работ сделана хорошо, выяснил, что есть еще огромное поле деятельности по улучшению большинства научных статей, которые он анализировал. В этом смысле работа ведется над двумя основными проблемами - воспроизводимость и повторяемость данных - чтобы справиться с этим кризисом.

2.2.4 Биостатистик в спорте: новая профессия, участвующая в профилактике травм

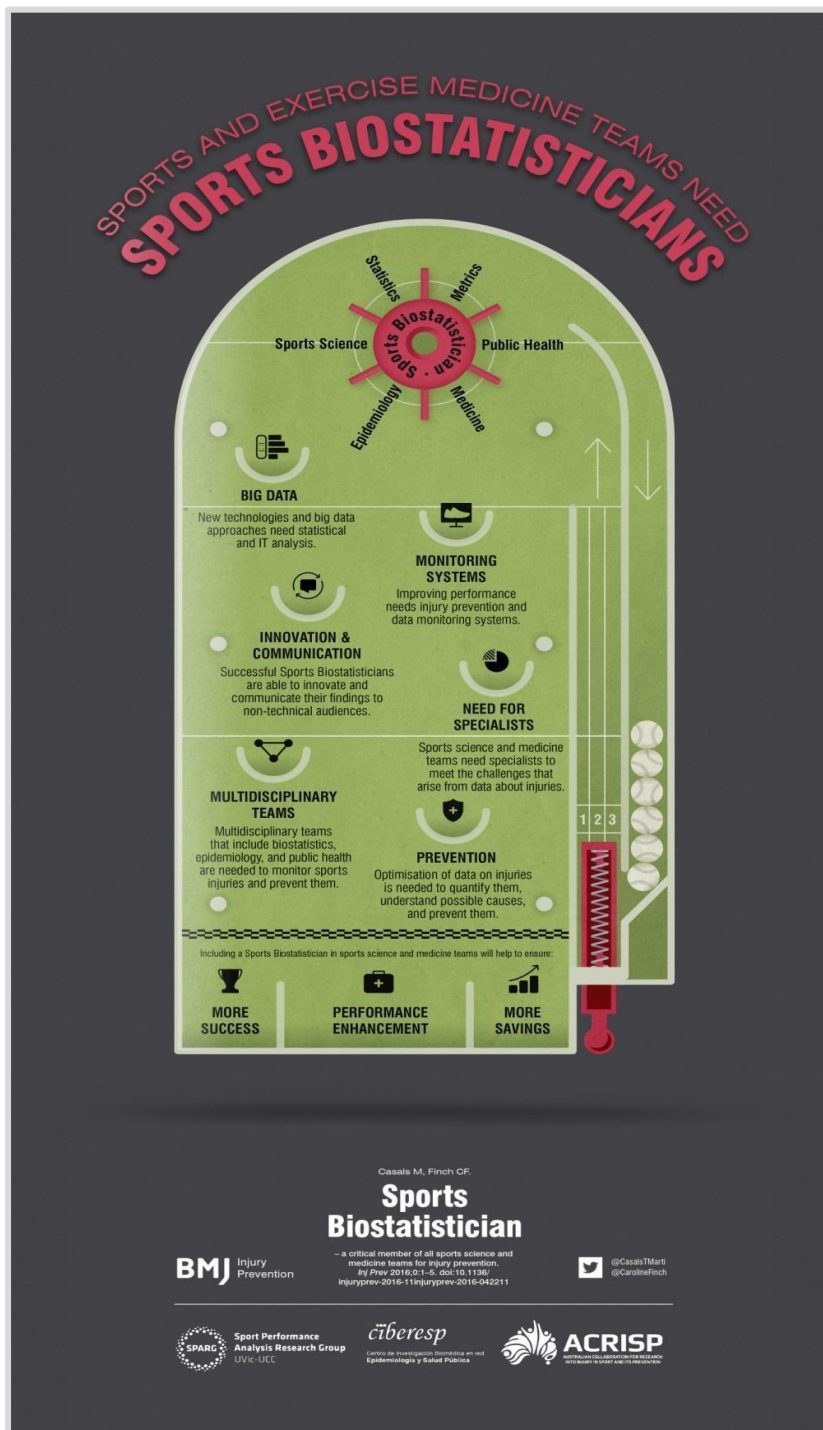
Новая профессиональная специальность спортивного биостатистика может помочь в оптимизации данных по травмам в виде их подсчёта, понимании их причин, а также наличии возможности вести профилактику” (“Sport Biostatistics Calls for Court to Prevent Injuries”, 2017, goo.gl/B633Kg), в соответствии с новым исследованием, опубликованным в Injury Prevention (Casals, and Finch, 2016). В этом исследовании Marti Casals, профессор и исследователь в Sport Performance Analysis Research Group (SPARG) из Университета Каталонии (University of Vic-Central University of Catalonia - UVic-UCC), в данный момент работает в роли биостатистика в ФК Барселона и с опытом полученным в команде НБА (Национальная Баскетбольная Ассоциация, США), вместе с Caroline Finch из Федерального Университета Австралии (Australian Collaboration for Research into Sports and its Prevention), являющимся одним из девяти центров исследований под эгидой МОК (Международный Олимпийский Комитет) по предотвращению травм и заботе о здоровье спортсменов, описывают поле деятельности спортивного биостатистика.

Говоря о спортивной производительности мы затрагиваем профилактику и предотвращение травм и систему наблюдения за ними.... Спортивная наука и медицина нуждаются в специалистах, отвечающих требованиям по обработке данных о травмах. Один из специалистов, способных делать эту работу - спортивный биостатистик, наиболее известна эта специальность в США или Австралии. (Casals, and Finch, cited in “La estadística también ayuda a prevenir lesiones”, 2016, goo.gl/UtpUHd).

Эта новая специальность требует знаний области этиологии травм, высокого уровня навыков в статистике, эпидемиологии, компьютерном программировании, как и сильные коммуникационные навыки, так как такому специалисту придется обсуждать свою работу с большим количеством разных людей - от родителей до тренеров, игроков, врачей, тренеров по ФП, физических терапевтов, спортивных ученых, эпидемиологов и тех, кто принимает решения в клубе (Pichel Andres, 2017).

В дополнение к спортивным аналитикам профессиональные клубы элитного уровня уже начали нанимать спортивных биостатистиков. Следующая инфографика, опубликованная в British Journal of Sport Medicine, показывает общие характеристики этой специальности (Casals, Bekker, and Finch, 2017).

Схема 2. Спортивный биостатистик



Источник: Casals et al., 2017.

Ссылки:

Barboza, S. D., Bolling, C. S., Nauta, J., Van Mechelen, M., y Verhagen, E. (2017). Приемлемость и восприятие конечными пользователями онлайн-системы наблюдения за спортивным здоровьем. *BMJ Open Sport & Exercise Medicine*, 3(1).

Casals, M., y Finch, C. F. (2016). Спортивный биостатистик – важнейший член всех спортивных научных и медицинских команд по профилактике травматизма. *Профилактика Травматизма*, 23(6), 423-427.

Casals, M., Bekker, S., y Finch, C. F. (2017). Инфографика: спортивные биостатистики как важнейший члены всех спортивных научных и медицинских команд по профилактике травматизма. *Inj Prev*. 2017 Dec;23(6):423-427.

Ioannidis, J. P. (2005). Почему большинство опубликованных результатов исследований являются ложными. *PLoS medicine*, 2(8).

Спортивная биостатистика призывает суд избегать травм. (3 января 2017 г.). Получено с <http://www.immedicohospitalario.es/noticia/10183/la-biestadistica-deportiva-pide-cancha-para-evitar-lesiones>

Статистика также помогает предотвратить травмы. (5 января 2017 г.). Синк, наука - это новости [Цифровая версия]. Получено с <http://www.agenciasinc.es/Noticias/La-estadistica-tambien-ayuda-a-prevenir-lesiones>.

Mabrouki, K., y Bosch, F. (2007). Научное письмо по биомедицине. Что тебе нужно знать. Барселона, ES: Prous Science.

Nielsen, R. O., Debes-Kristensen, K., Hulme, A., Bertelsen, M. L., Møller, M., Parner, E. T., y Mansournia, M. A. (2017A). Лучше ли показатели распространенности, чем показатели заболеваемости, в исследованиях спортивных травм? *Br J Sports Med*. 2017 Октябрь 13.

Nielsen, R. O., Chapman, C. M., Louis, W. R., Stovitz, S. D., Mansournia, M. A., Windt, J., Møller, M., Thorlund Parner, E., Hulme, A., Lejbach Bertelsen, M., Finch, C. F., Casals, M., y Verhagen, E. (2017b). Семь грехов при интерпретации статистики в науке о спортивном травматизме. *Br J Sports Med*. 2017 20 декабря.

Nielsen, R. O., Bertelsen, M. L., Verhagen, E., Mansournia, M. A., Hulme, A., Møller, M., y Casals, M. (2017). Когда результат исследования важен для спортсменов, врачей и тренеров / персонала команд? *Британский журнал спортивной медицины*, 51, 1454-1455.

Pichel Andrés, J. (16 de enero de 2017). Биостатистика и «большие данные» для предотвращения спортивных травм. Получено с <https://www.bez.es/882750780/bioestadistica-big-data-para-prevenir-lesiones.html>.

Американская статистическая ассоциация (ASA). (7 марта 2016 г.). Американская статистическая ассоциация опубликовала заявление о статистической значимости и P-значениях. Предлагает принципы для улучшения проведения и интерпретации количественной науки. Получено с <http://www.amstat.org/asa/files/pdfs/P-ValueStatement.pdf>.

Verhagen, E., Stovitz, S. D., Mansournia, M. A., Nielsen, R. O., y Shrier, I. (2017). Образовательные передовые статьи BJSM: важны методы. Br J Sports Med. 2017 авг. 17.