

СОВРЕМЕННЫЕ КЛАССИФИКАЦИИ ПОВРЕЖДЕНИЙ НИЖНЕШЕЙНОГО ОТДЕЛА ПОЗВОНОЧНИКА. ЧАСТЬ 2. СИСТЕМАТИЧЕСКИЙ ОБЗОР ИССЛЕДОВАНИЙ НАДЕЖНОСТИ СУЩЕСТВУЮЩИХ ШКАЛ И ВОСПРОИЗВОДИМОСТИ РЕЗУЛЬТАТОВ ИХ ПРИМЕНЕНИЯ¹

А.А. Гринь^{1,2}, И.С. Львов¹, С.Л. Аракелян³, А.Э. Тальпов¹, А.Ю. Кордонский¹, А.В. Сытник³,
Б.А. Абдухаликов¹, У.Г. Хушназаров¹, В.А. Каранадзе^{1,2}, В.В. Крылов^{1,2}

¹ГБУЗ «Научно-исследовательский институт скорой помощи им. Н.В. Склифосовского Департамента здравоохранения г. Москвы»; Россия, 129090 Москва, Большая Сухаревская пл., 3;

²ФГБОУ ВО «Московский государственный медико-стоматологический университет им. А.И. Евдокимова» Минздрава России; Россия, 127473 Москва, ул. Делегатская, 20, стр. 1;

³ГБУЗ «Городская клиническая больница №13 Департамента здравоохранения г. Москвы»; Россия, 115280 Москва, ул. Велозаводская, 1/1

Контакты: Антон Юрьевич Кордонский akord.neuro@mail.ru

Цель исследования — провести систематический обзор зарубежных и отечественных исследований и определить оптимальную для использования в клинической практике классификацию повреждений нижнешейного отдела позвоночника.

Материалы и методы. Систематический обзор выполнен в соответствии с рекомендациями PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Поиск проводили среди англоязычных статей, представленных в базе данных PubMed, и среди русскоязычных статей, размещенных на платформе eLIBRARY.ru. Критерии включения в обзор: наличие полнотекстовой версии статьи на английском или русском языке; возраст пациентов ≥ 18 лет; наличие информации об одной из фаз валидации классификаций по рекомендациям L. Audige и соавт.

Результаты. Для изучения полного текста были отобраны 30 статей, из них все работы на русском языке ($n = 3$) были опубликованы одним авторским коллективом, данные дублировались и не содержали необходимых статистических показателей, в связи с чем они были исключены из исследования. Из 27 статей на английском языке 8 соответствовали критериям и были включены в обзор. Согласно результатам их анализа, наибольшей надежностью и воспроизводимостью результатов характеризуются классификации AOSpine и Subaxial Injury Classification System. Шкала Allen–Fergusson имеет более низкий коэффициент согласия экспертов, но при этом она способна дать более наглядное представление о повреждении. Так же поверхностно изучена классификация J. Harris и соавт. Надежность шкалы C. Argenson и соавт. не исследована. Данных для полноценной оценки Cervical Spine Injury Severity Score имеющиеся публикации не представляют. Анализ работ наглядно продемонстрировал необходимость более тщательного исследования всех существующих шкал и классификаций. Данное исследование должно быть проведено на базе нескольких стационаров участниками с разным уровнем опыта (от резидента до опытного спинального хирурга). Должна быть проанализирована не только воспроизводимость результатов применения отдельных классификаций, но и параметр обучаемости и взаимосвязь отдельных шкал и классификаций.

Ограничения исследования обусловлены недостаточным количеством публикаций, небольшим объемом выборок и неоднородностью групп, различием опыта экспертов.

Заключение. Наиболее надежными из представленных являются классификации AOSpine и Subaxial Injury Classification System. Тем не менее имеющиеся в литературе данные недостаточно для полноценного сравнения шкал и классификаций. Необходимы дальнейшие многоцентровые исследования надежности классификаций для выбора среди них оптимальной.

Ключевые слова: повреждения нижнешейного отдела позвоночника, оценка, Subaxial Injury Classification System, Cervical Spine Injury Severity Score, SLIC, CSISS, классификация Allen–Fergusson, классификация J. Harris и соавт., классификация C. Argenson и соавт., классификация AOSpine

Для цитирования: Гринь А.А., Львов И.С., Аракелян С.Л. и др. Современные классификации повреждений нижнешейного отдела позвоночника. Часть 2. Систематический обзор исследований надежности существующих шкал и воспроизводимости результатов их применения. *Нейрохирургия* 2019;21(2):28–38.

DOI: 10.17650/1683-3295-2019-21-2-28-38

¹Первая часть обзора опубликована в № 1 журнала «Нейрохирургия» за 2019 г.

Currently available classification systems for lower cervical spine injuries. Part 2. Systematic review of studies on the reliability and reproducibility of existing scales

A.A. Grin^{1,2}, I.S. Lvov¹, S.L. Arakelyan³, A.E. Talypov¹, A. Yu. Kordonsky¹, A.V. Sytnik³, B.A. Abdukhalikov¹,
U.G. Khushnazarov¹, V.A. Karanadze^{1,2}, V.V. Krylov^{1,2}

¹N.V. Sklifosovsky Research Institute for Emergency Medicine, Moscow Healthcare Department;
3 Bol'shaya Sukharevskaya Sq., Moscow 129090, Russia;

²A.I. Evdokimov Moscow State University of Medicine and Dentistry, Ministry of Health of Russia; Bld. 1, 20 Delegatskaya St.,
Moscow 127473, Russia;

³City Clinical Hospital No. 13, Moscow Healthcare Department; 1/1 Velozavodskaya St., Moscow 115280, Russia

The study objective is to review the Russian and foreign studies and to identify an optimal classification system for lower cervical spine injuries. **Materials and methods.** This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). We conducted a search for articles published in English (PubMed database) and Russian (eLIBRARY.ru). The inclusion criteria were as follows: available full text, patient age ≥ 18 years, and information on one of the validation phases for classifications according to L. Audige et al.

Results. A total of 30 articles were eligible. Of them, 3 studies were published in Russian (by one group of authors); however, they didn't contain required statistical parameters and had duplicated data; therefore, they were excluded from the analysis. Out of 27 articles published in English, 8 articles met all the criteria and were included into the systematic review. The AOSpine and Subaxial Injury Classification Systems demonstrated the highest reliability and reproducibility of the results. The Allen–Fergusson classification has lower intraobserver and interobserver agreement coefficients, but it can give a clearer visual representation of injuries. We also assessed J. Harris classification system. The reliability of the scale developed by C. Argenson et al. was not evaluated. The analyzed publications contained no data for full evaluation of the Cervical Spine Injury Severity Score. Our analysis clearly demonstrated the need for a more thorough evaluation of all available scales and classifications. This study should be multicenter and involve experts with different levels of experience (from residents to experienced spinal surgeons). Moreover, it should analyze not only the reproducibility of individual classifications, but also the aspects of learning and the relationship between individual scales and systems.

The main study limitations included insufficient number of publications, small sample sizes, heterogeneity of groups, and differences in the experience of experts.

Conclusion. The AOSpine and Subaxial Injury Classification Systems are the most reliable classification systems. However, the data available in literature is not sufficient for a full comparison of all existing scales and systems. Further multicenter studies on the reliability of classifications are needed to select an optimal one.

Key words: lower cervical spine injuries, scoring systems, Subaxial Injury Classification System, Cervical Spine Injury Severity Score, SLIC, CSISS, Allen–Fergusson classifications, J. Harris et al. classification, C. Argenson et al. classification, AOSpine classification

For citation: Grin A.A., Lvov I.S., Arakelyan S.L. et al. Currently available classification systems for lower cervical spine injuries. Part 2. Systematic review of studies on the reliability and reproducibility of existing scales. *Neyrokhirurgiya = Russian Journal of Neurosurgery* 2019;21(2):28–38.

ВВЕДЕНИЕ

За последние 40 лет в общемировую практику внедрены многие классификации повреждений нижнешейного отдела позвоночника (НШОП). Наиболее популярными из них были морфологические классификации F. Holdsworth, Allen–Fergusson, J. Harris и соавт., C. Argenson и соавт., AOSpine, а также шкалы балльной оценки повреждений White & Panjabi, Subaxial Cervical Spine Injury Classification System (SLIC) и Cervical Spine Injury Severity Score (CSISS) [1]. Несмотря на то что данные шкалы применяют достаточно давно, повсеместного распространения ни одна из них так и не получила. Согласно результатам крупнейшего анкетирования, проведенного H.S. Chhabra и соавт. [2], только 37,5 % экспертов в ежедневной практике для оценки повреждений НШОП применяли классификацию Allen–Fergusson, 35 % – SLIC, 5 % – CSISS, 7,5 % – другие шкалы, а 15 % респондентов в рутинной практике не использовали классификации вообще. Большинство опрошенных отметили, что да-

же выбранная ими классификация недостаточно удобна для ежедневного применения и не помогает составить эффективный лечебный алгоритм.

Валидация любой классификации или шкалы включает 3 фазы [3]. В ходе 1-й фазы проводят пилотное исследование в 1 стационаре, после чего, в случае необходимости, классификацию (шкалу) дорабатывают. Вторая фаза предусматривает многоцентровое исследование с привлечением хирургов с разным уровнем опыта. В 3-й фазе классификацию (шкалу) внедряют в клиническую практику, чтобы сопоставить результаты ее применения с исходами заболевания.

Публикации, посвященные валидации классификаций повреждений НШОП, носят единичный характер и, как правило, представляют результаты 1-й фазы валидации.

Цель данного исследования – провести систематический анализ зарубежных и отечественных исследований и определить оптимальную для использования в клинической практике классификацию повреждений НШОП.

МАТЕРИАЛЫ И МЕТОДЫ

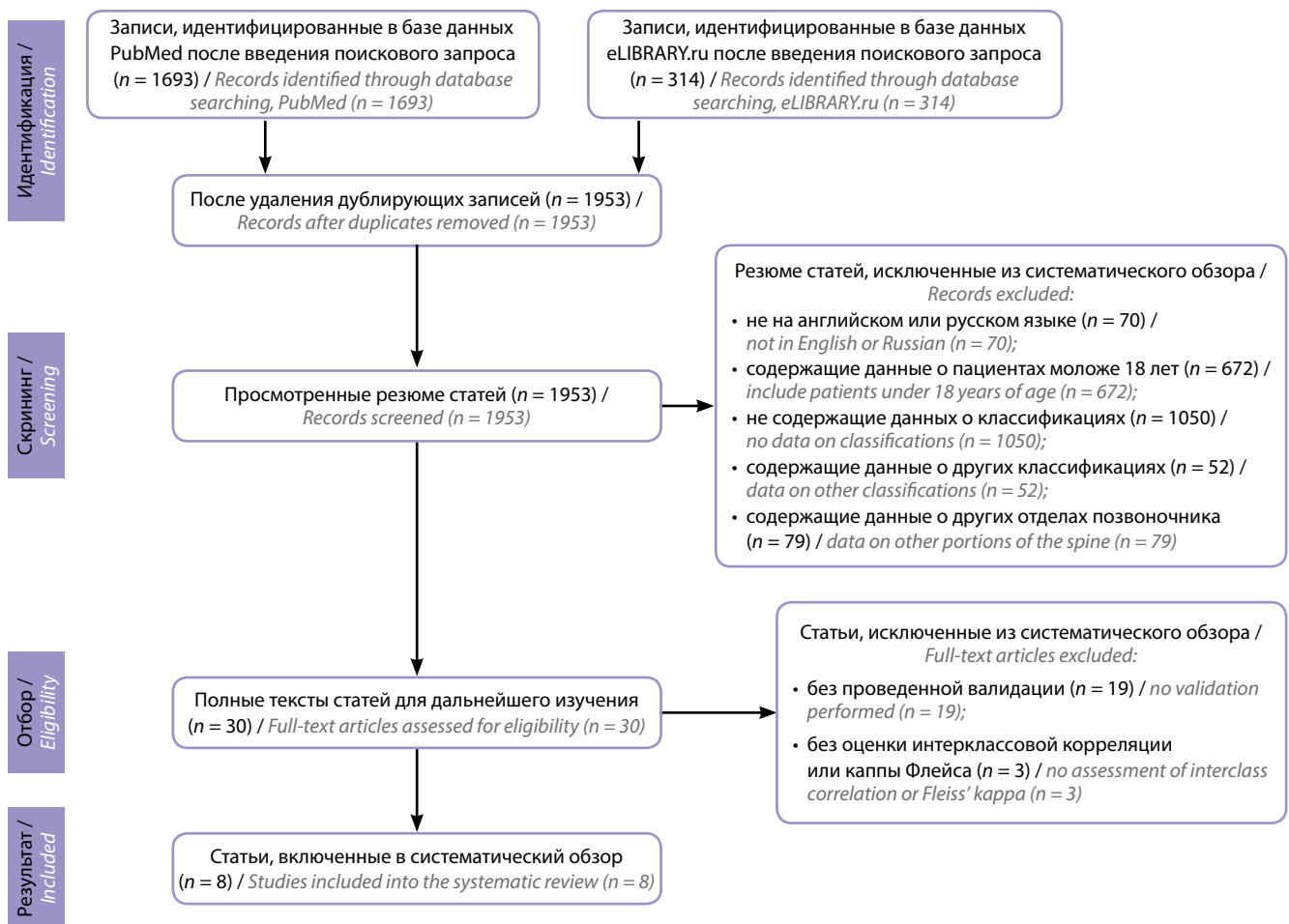
Отбор статей. Систематический обзор выполнен в соответствии с рекомендациями, отраженными в Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [4]. Поиск в базе данных PubMed был выполнен среди англоязычных статей при помощи ключевых слов и базовых операторов: (“subaxial” OR “cervical” OR “Allen” OR “Fergusson” OR “Harris” OR “AOSpine” OR “Argenson” OR “SLIC” OR “CSISS”) AND (“system” OR “classification” OR “score” OR “reliability” OR “validation” OR “assessment” OR “teachability”) NOT (“lumbar” OR “thoracic” OR “thoracolumbar” OR “hip” OR “ankle” OR “screw”). Для поиска русскоязычных статей, размещенных на платформе eLIBRARY.ru, применяли ключевые слова и базовые операторы: («субаксиальный» OR «шейный» OR «нижнешейный») AND («шкала» OR «классификация»).

Критерии включения в исследование: 1) наличие полного текста статьи на английском или русском языке; 2) возраст пациентов, данные которых представлены в статье, ≥ 18 лет; 3) описание одной из фаз валидации классификаций в соответствии с рекомен-

дациями L. Audige и соавт. [3]. Все статьи, не соответствовавшие данным критериям, были исключены из обзора. Алгоритм поиска и отбора статей представлен на рисунке.

Формирование базы данных. Полученные из каждой статьи данные заносили в программу Microsoft Excel (Office 2016 for Mac), распределяя их по ячейкам в соответствии с типом информации: годы проведения исследования, число пациентов, виды диагностических исследований, которые применялись в процессе валидации (рентгенография, компьютерная (КТ) и магнитно-резонансная (МРТ) томография или их комбинация), общее число экспертов, уровень опыта хирургов и их специализация, названия шкал, количество центров, участвовавших в исследовании. Результаты валидации каждой из шкал заносили в ячейки в соответствии с каждым типом и подтипом. Если какая-либо информация в тексте статьи отсутствовала, соответствующую ячейку помечали как «н. д.» («нет данных»).

Статистический анализ. Статистический анализ осуществляли с помощью программы Microsoft Excel



Алгоритм поиска и отбора статей в соответствии с Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA)

Algorithm of search and selection of articles according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA)

(Office 2016 for Mac). Сравнение проводили методами описательной статистики. Для морфологических классификаций сравнивали коэффициенты согласованности заключений одного и того же исследователя (intra-observer reliability) и заключений разных исследователей (interobserver reliability) с применением каппы Флейса (Fleiss kappa). Для шкал балльной оценки анализировали также коэффициент внутрикласовой корреляции (intraclass correlation coefficient, ICC). Результаты интерпретировали в соответствии с системой Landis & Koch [5]. Если каппа или ICC не превышали 0,2, то степень согласованности считали слабой, при значениях от 0,2 до 0,4 – удовлетворительной, от 0,4 до 0,6 – умеренной, от 0,6 до 0,8 – значительной, >0,8 – превосходной.

РЕЗУЛЬТАТЫ

Отбор статей. При первоначальном поиске в PubMed выявили 1953 резюме статей. После фильтрации по возрасту пациентов и языку оставшиеся статьи были просмотрены. В результате было отобрано 30 публикаций для изучения полнотекстовых версий: 3 на русском языке и 27 на английском. Все работы на русском языке были написаны одним авторским коллективом; данные в них дублировались и не содержали необходимых статистических показателей (каппы Флейса или ICC), в связи с чем они были исключены из дальнейшего рассмотрения [6–8]. Из 27 статей на английском языке 8 соответствовали необходимым критериям и были включены в настоящее исследование.

Общая характеристика отобранных статей. Всего мы отобрали 8 статей [9–16]. В 5 работах исследователи описали результаты валидации только какой-либо 1 шкалы/классификации, в 3 статьях – 2 или 3 шкалы/классификаций. Число пациентов, включенных в эти исследования, варьировало от 11 до 65 человек, медиана – 50. Источником данных о состоянии НШОП были преимущественно изображения, полученные при КТ. В 4 работах данные КТ при необходимости дополняли результатами МРТ. В 1 статье источник данных указан не был [14]. В большинстве (6 (75 %)) исследований наиболее репрезентативные изображения КТ и МРТ отбирали соответствующие специалисты, которые впоследствии не участвовали в исследовании в качестве экспертов. В 2 (25 %) исследованиях участвовали все пациенты с травмой НШОП, проходившие лечение в определенный промежуток времени. Число экспертов варьировало от 5 до 20 человек, медиана – 11. Основную массу экспертов составляли врачи, специализирующиеся на спинальной хирургии или травматологии и ортопедии. Только в 3 исследованиях принимали участие резиденты. Большинство (5 (62,5 %) исследований проводились в пределах 1 клиники, а 3 (37,5 %) были многоцентровыми. Общая характеристика работ, посвященных валидации шкал оценки повреждений НШОП, представлена в табл. 1. После-

дующий подробный анализ показателей согласованности заключений экспертов представлен в табл. 2 и 3.

Классификация Allen–Fergusson. Данная классификация была исследована в 3 статьях [12–14]. Медиана каппы, отражающей согласованность заключений разных исследователей, для всех 6 стадий классификации составила 0,5 (умеренная согласованность). Постадийно каппа была рассчитана только в 2 статьях [12, 13]. Наилучшие показатели (умеренная или значительная степень согласованности) были выявлены в отношении повреждений, формирующихся вследствие вертикальной компрессии (vertical compression injuries) ($K = 0,61$ и $0,58$), дистракционно-флексионных (distractive flexion injuries) ($K = 0,54$ и $0,57$) и дистракционно-экстензионных (distractive extension injuries) ($K = 0,63$ и $0,40$). Слабой была степень согласованности заключений о латерофлексионных повреждениях (lateral flexion injuries) ($K = -0,16$ и $0,30$). Медиана каппы, отражающей согласованность заключений одного и того же исследователя, составила 0,66 (значительная степень согласованности). Только в 1 статье [13] каппа, отражающая согласованность заключений одного и того же исследователя, была рассчитана для каждого типа повреждений. Высокие степени согласованности также были достигнуты в отношении повреждений, формирующихся вследствие вертикальной компрессии ($K = 0,71$), дистракционно-флексионных ($K = 0,69$) и дистракционно-экстензионных ($K = 0,72$).

Классификация J. Harris и соавт. Надежность данной классификации и воспроизводимость результатов ее применения исследована только в 1 работе [14]. А. R. Vassago и соавт. представили коэффициенты согласованности заключений одного и того же исследователя и заключений разных исследователей – каппы Флейса – только относительно всех групп повреждений в целом без подробного анализа по стадиям. Степень согласованности заключений одного и того же исследователя была умеренной ($K = 0,53$). Степень согласованности заключений разных исследователей была существенно ниже и лишь приближалась к удовлетворительной ($K = 0,41$).

Классификация AOSpine. Данная классификация исследована более подробно в сравнении с прочими. Результаты валидации AOSpine представлены в 3 работах, опубликованных в 2016 г. [11, 13, 15]. Медиана коэффициента согласованности заключений разных исследователей – каппы Флейса – составила 0,63 (0,61–0,65), т. е. уровень согласованности был значительным. При подробном анализе по типам повреждения медиана каппы в отношении компрессионных повреждений (типа А) составила 0,66 (0,64–0,67), в отношении дистракционных (типа В) – 0,51 (0,08–0,54), в отношении трансляции позвонков (типа С) – 0,68 (0,65–0,73). В работе О. Т. Silva и соавт. каппа Флейса, отражающая согласованность заключений разных исследователей в отношении повреждений суставных

отростков (типа F), рассчитывалась нами (в виде арифметического среднего), поскольку авторы указали значения этого коэффициента только для подтипов данного повреждения; медиана каппы составила 0,61 (0,51–0,66). Уровень согласованности заключений одного и того же эксперта был превосходным, для всех типов повреждений медиана каппы Флейса составила 0,77 (0,68–0,90) [11]. Подробный расчет коэффициентов для разных типов повреждений был выполнен только в 1 работе [13]. Для повреждений типа А он составил 0,73, для типа В – 0,65, С – 0,70, F – 0,61.

Шкала SLIC. Данные о валидации этой шкалы опубликованы в 3 работах [10, 12, 14]. Медиана коэффициента согласованности заключений разных исследователей – ICC – составила 0,78 (значительный уровень), заключений одного и того же исследователя – 0,90 (превосходный уровень). Коэффициент согласованности заключений разных экспертов (каппа Флейса) и доля совпадений в отношении морфологии повреждения, целостности диско-связочного комплекса и неврологического статуса были оценены только в 2 работах [10, 14]. При оценке морфологии повреждения средняя доля совпадений составила 55,4 % (47,8–63,0 %), средний коэффициент согласованности заключений разных экспертов (каппа Флейса) – 0,40 (0,29–0,51). При оценке целостности диско-лигаментарного комплекса средняя доля совпадений была больше и составила 62,5 % (58,0–67,1 %), каппа Флейса – 0,40 (0,46–0,33). При оценке неврологического статуса средняя доля совпадений составила 79,3 % (87,6–71,0 %), средняя каппа Флейса – 0,66 (0,70–0,62). Умеренный уровень согласованности был достигнут в отношении хирургической тактики [10, 14] – 74,1 % (K = 0,55) и 73,9 % (K = 0,44), что соответствовало средним значениям ICC и каппы Флейса. Тем не менее выбранная при помощи шкалы тактика соответствовала личным предпочтениям хирургов в 93,3 % случаев [14].

Шкала CSISS. Данные о валидации шкалы CSISS представлены в 3 публикациях [9, 12, 16]. Эта шкала имеет наиболее высокие показатели ICC. Медиана ICC, отражающего согласованность заключений одного эксперта, составила 0,980 (0,977–0,983), разных экспертов – 0,960 (0,883–0,975). Каппа Флейса для шкалы CSISS ни в одной работе не была рассчитана.

Сопоставление классификаций. Между собой 2 разные шкалы были сопоставлены только в 1 работе [14]. Авторы сопоставили заключения экспертов, касающиеся оценки морфологии перелома по шкале SLIC и соответствующие типы повреждений по классификации Allen–Fergusson. Они совпали на 71,5 % (K = 0,61).

ОБСУЖДЕНИЕ

Вопрос о создании идеальной классификации повреждений позвоночника остается открытым долгое время. Согласно критериям S.K. Mirza, такая шкала

должна: 1) позволить идентифицировать и описать любое повреждение; 2) отразить патогенез и механизм повреждения, обозначить степень повреждения; 3) задать направление для выбора метода лечения; 4) содержать легко распознаваемые клинические и радиологические характеристики; 5) отражать степень неврологических нарушений; 6) содержать данные о степени костных и связочных повреждений; 7) предсказывать результаты лечения, определять риск развития деформации, дополнительного повреждения невральных структур; 8) подготовить почву для последующих исследований [17]. В 2009 г. J. J. van Middendorp и соавт. также описали свойства идеальной классификации. По их мнению, она должна: 1) содержать строгие определения, не предполагающие двусмысленной или свободной трактовки; 2) быть всеобъемлющей и состоять из взаимоисключающих категорий; 3) иметь ясные графические иллюстрации; 4) быть простой и удобной для ежедневного применения; 5) включать ограниченное количество категорий; 6) отражать градацию (повышение) степени тяжести повреждений; 7) иметь буквенно-цифровые обозначения для каждой группы и подгруппы; 8) описывать четкие отличительные признаки каждого типа повреждений на диагностических снимках [18].

К сожалению, идеальной классификации к настоящему моменту разработать не удалось. Если классификация описывает все механизмы повреждений и предполагает составление хирургического алгоритма, то она слишком громоздка и сложна для запоминания (имеет пологую кривую обучаемости). Если результаты применения классификации хорошо воспроизводимы, то выделенные в ней типы и подтипы не дают четкого представления о самом механизме повреждения. Исследования надежности существующих классификаций немногочисленны и содержат противоречивые данные.

Так, при проведении систематического обзора мы выявили, что анализу наиболее известной классификации повреждений НШОП (Allen–Fergusson) были посвящены всего 3 работы. Данная шкала включает достаточно большое количество подтипов повреждений и рассматривает почти все возможные механизмы переломов НШОП. Это же и является основным ее недостатком. Наличие 6 типов повреждений обусловило умеренный уровень согласованности заключений экспертов между собой и значительный уровень согласованности повторных заключений. Доля совпадений при этом составила 71,4 % [15]. Было проведено только 1 многоцентровое исследование, при этом все эксперты были опытными спинальными хирургами или ортопедами.

Классификация J. Nagris и соавт. исследована только в 1 работе [14]. К сожалению, предоставленных в статье данных недостаточно для того, чтобы составить суждение об этой шкале. Несмотря на то что типов повреждения в ней описано больше (7), общее

Таблица 1. Общая характеристика исследований, посвященных валидации классификаций повреждений нижнейшейного отдела позвоночника и шкал их оценки
Table 1. Overall characteristics of studies evaluating reliability of classification systems for subaxial injuries

Авторы Authors	Год публикации Year of publication	Классификация (шкала) Classification (scale)	Число пациентов Number of patients	Формат анализируемых данных Data analyzed	Метод формирования выборки пациентов Patient sampling	Число экспертов Number of experts	Эксперты Experts		Тип исследования Type of study
							Специализация Specialty	Число Number	
P.A. Anderson и соавт. [9] P.A. Anderson et al. [9]	2007	CSISS	34	Рентгенограмма, КТ X-ray, CT	Отбор снимков Selected images	15	Резиденты Residents	Н. д. N. d.	Моноцентровое Single-center
							Лечащие врачи Attending physicians	Н. д. N. d.	
							Спинальные хирурги Spinal surgeons	Н. д. N. d.	
S.W. Zehnder и соавт. [16] S.W. Zehnder et al. [16]	2009	CSISS	50	Рентгенограмма, КТ X-ray, CT	Отбор снимков Selected images	15	Рентгенолог Radiologist	1	Моноцентровое Single-center
							Младшие резиденты Junior residents	6	
							Старшие резиденты Senior residents	6	
J. J. van Middendorp и соавт. [10] J. J. van Middendorp et al. [10]	2013	SLIC	51	КТ, МРТ CT, MRI	Отбор снимков Selected images	12	Спинальные хирурги Spinal surgeons	3	Моноцентровое Single-center
							Нейрохирурги Neurosurgeons	Н. д. N. d.	
							Травматологи-ортопеды Traumatologists/orthopedists	Н. д. N. d.	
A.T. Stone и соавт. [12] A.T. Stone et al. [12]	2010	SLIC, CSISS, Allen-Ferguson	50	КТ, МРТ CT, MRI	Сплошная выборка за временной интервал Consecutive images over a time period	5	Спинальные хирурги Spinal surgeons	5	Моноцентровое Single-center
							Травматологи-ортопеды Traumatologists/orthopedists	5	
A.R. Vaccaro и соавт. [14] A.R. Vaccaro et al. [14]	2007	SLIC, Allen-Ferguson, J. Harris	11	Н. д. N. d.	Отбор снимков Selected images	20	Спинальные хирурги Spinal surgeons	5	Многоцентровое Multicenter
							Травматологи-ортопеды Traumatologists/orthopedists	15	
J. Urrutia и соавт. [13] J. Urrutia et al. [13]	2016	AOSpine, Allen-Ferguson	65	КТ, МРТ CT, MRI	Отбор снимков Selected images	6	Спинальные хирурги Spinal surgeons	3	Моноцентровое Single-center
							Травматологи-ортопеды Traumatologists/orthopedists	3	
A.R. Vaccaro и соавт. [15] A.R. Vaccaro et al. [15]	2015	AOSpine	30	КТ, МРТ CT, MRI	Отбор снимков Selected images	10	Спинальные хирурги Spinal surgeons	10	Многоцентровое Multicenter

Окончание табл. 1
The end of the table 1

Авторы Authors	Год публикации Year of publication	Классификация (шкала) Classification (scale)	Число пациентов Number of patients	Формат анализированных данных Data analyzed	Метод формирования выборки пациентов Patient sampling	Число экспертов Number of experts	Эксперты Experts		Тип исследования Type of study
							Специализация Specialty	Число Number	
O. T. Silva и соавт. [11] O. T. Silva et al. [11]	2016	AOSpine	51	КТ CT	Сплошная выборка за временной интервал Consecutive images over a time period	5	Младший резидент Junior resident	1	Моноцентровое Single-center
							Старшие резиденты Senior residents	2	
							Нейрохирург Neurosurgeon	1	
							Травматолог-ортопед Traumatologist/orthopedist	1	

Примечание. КТ – компьютерная томография; МРТ – магнитно-резонансная томография; н. д. – нет данных; SLIC – Subaxial Cervical Spine Injury Classification System; CSISS – Cervical Spine Injury Severity Score.
Note. CT – computed tomography; MRI – magnetic resonance imaging; n. d. – no data; SLIC – Subaxial Cervical Spine Injury Classification System; CSISS – Cervical Spine Injury Severity Score.

Таблица 2. Коэффициенты согласованности заключений одного и того же исследователя и заключений разных исследователей (каппы Флейса) при валидации классификаций Allen–Ferguson, J. Harris и соавт. и AOSpine

Table 2. Intraobserver and interobserver reliability coefficients (Fleiss' kappa) in the Allen–Ferguson, J. Harris et al., and AOSpine studies

Авторы Authors	Год публикации Year of publication	Коэффициент согласованности заключений одного и того же исследователя Intraobserver reliability coefficient			Коэффициент согласованности заключений разных исследователей Interobserver reliability coefficient						
		Allen–Ferguson	J. Harris и соавт. J. Harris et al.	AOSpine	Allen–Ferguson	J. Harris и соавт. J. Harris et al.	AOSpine				
							Общая оценка Total score	A	B	C	F
O. T. Silva и соавт. [11] O. T. Silva et al. [11]	2016	–	–	0,9	–	–	0,63	0,67	0,08	0,68	0,51
A. T. Stone и соавт. [12] A. T. Stone et al. [12]	2010	0,91	–	–	0,50	–	–	–	–	–	–
J. Urrutia и соавт. [13] J. Urrutia et al. [13]	2016	0,66	–	0,68	0,46	–	0,61	0,64	0,51	0,65	0,61
A. R. Vaccaro и соавт. [14] A. R. Vaccaro et al. [14]	2007	0,63	0,53	–	0,53	0,41	–	–	–	–	–
A. R. Vaccaro и соавт. [15] A. R. Vaccaro et al. [15]	2015	–	–	0,77	–	–	0,65	0,66	0,54	0,73	0,66
Медиана каппы Флейса Fleiss' kappa median		0,66	0,53	0,77	0,50	0,41	0,63	0,66	0,51	0,68	0,61

Таблица 3. Коэффициенты согласованности заключений одного и того же исследователя и заключений разных исследователей (коэффициенты внутрикласовой корреляции) при валидации шкал Subaxial Cervical Spine Injury Classification System и Cervical Spine Injury Severity Score
Table 3. Intraobserver and interobserver reliability coefficients (intraclass correlation coefficients) in Subaxial Cervical Spine Injury Classification System and Cervical Spine Injury Severity Score studies

Авторы Authors	Год публикации Year of publication	Коэффициент согласованности заключений одного и того же исследователя Intraobserver reliability coefficient		Коэффициент согласованности заключений разных исследователей Interobserver reliability coefficient	
		SLIC	CSISS	SLIC	CSISS
P.A. Anderson и соавт. P.A. Anderson et al.	2007	–	0,977	–	0,883
J. J. van Middendorp и соавт. [10] J. J. van Middendorp et al. [10]	2013	–	–	0,780	–
A.T. Stone и соавт. [12] A.T. Stone et al. [12]	2010	0,980	0,980	0,790	0,960
A.R. Vaccaro и соавт. [14] A.R. Vaccaro et al. [14]	2007	0,830	–	0,710	–
S.W. Zehnder и соавт. [16] S.W. Zehnder et al. [16]	2009	–	0,983	–	0,975
Медиана коэффициента внутрикласовой корреляции Intraclass correlation coefficient median		0,900	0,980	0,780	0,960

Примечание (Note). SLIC – Subaxial Cervical Spine Injury Classification System; CSISS – Cervical Spine Injury Severity Score.

количество подтипов меньше, чем в классификации Allen–Fergusson (12 и 23 соответственно). Меньшее количество подтипов было выделено для дистракционных переломов и вывихов, и в отдельный тип объединены ротационные повреждения. Описана значительная часть переломов НШОП. Все это, казалось бы, должно было упростить классифицирование повреждений НШОП, тем не менее каппа Флейса, отражающая согласованность заключений экспертов, у этой классификации была самой низкой. Анализ классификации по стадиям не проводили, в связи с чем судить о ее конкретных достоинствах и недостатках затруднительно. Несмотря на то что работа A.R. Vaccaro и соавт. была многоцентровой, ее существенным недостатком является малый размер выборки больных. С учетом вышеизложенного объективно судить о надежности данной классификации и воспроизводимости результатов ее использования по данным литературы не представляется возможным.

Валидация классификации AOSpine была описана в 3 статьях. По сравнению с другими описательными классификациями она существенно упрощена – до 3 типов и 9 подтипов повреждений. Дополнительно выделены типы повреждений дугоотростчатых суставов для уточнения хирургической тактики. Результатом упрощения классификации стал значительный уровень согласованности заключений (как разных экспертов, так и одного и того же эксперта). Возможно, это обуславливает и тот факт, что большинство экспертов (85,7 %) были опытными специалистами

в области спинальной хирургии, а 2 из 3 работ были моноцентровыми. При анализе по типам повреждений существенно различались заключения в отношении повреждений типов В и F только в работе O.T. Silva и соавт. [11]. На наш взгляд, причина этого в том, что дистракционные повреждения без трансляции позвонков встречаются достаточно редко, а пациенты в этой работе отбирались методом сплошной выборки за определенный промежуток времени. В 2 других исследованиях [13, 15] проводили специальный отбор снимков в соответствии с каждым типом повреждения. O.T. Silva и соавт. отмечают также значительно более низкие значения каппы Флейса при первичной оценке ($\leq 0,53$) (удовлетворительный уровень согласованности), в то время как при повторной оценке он превышал 0,60 (значительный уровень). Этот факт свидетельствует о формировании определенной кривой обучаемости, тем более что авторы указывают на отсутствие опыта применения этой шкалы у 2 участников. Работа O.T. Silva и соавт. – это единственное исследование классификации AOSpine, в котором резиденты принимали участие наравне с опытными спинальными хирургами. К сожалению, параметр обучаемости в данной статье не обсуждается.

Шкала SLIC – первая система балльной оценки, учитывающая морфологию повреждения, целостность дисколигаментарного комплекса и степень неврологических нарушений. Она была разработана для упрощения принятия решения о необходимости хирургического лечения. Уровень внутрикласовой корреляции во всех

3 работах был значительным и приближался к превосходному. Но несмотря на достаточно высокий ИСС, в имеющихся статьях доля совпадений была достаточно маленькой в отношении конкретных числовых значений. На это указывает и низкое значение каппы Флейса (в среднем 0,4) при оценке морфологии повреждения и целостности дисколигаментарного комплекса. Возможной причиной такого расхождения при оценке морфологии повреждений, на наш взгляд, было участие разных нейрохирургических и травматологических школ в 2 многоцентровых работах [10, 14]. Работа, в которой все участники были из одного и того же учреждения [12], продемонстрировала наиболее высокие показатели ИСС, но каппу Флейса авторы статьи не указали, поэтому оценить точность совпадения заключений экспертов в этом исследовании не представляется возможным. Тем не менее во всех исследованиях была достаточно высокой доля совпадений заключений в отношении дальнейшей хирургической тактики.

Наиболее спорной из анализируемых была шкала CSISS. На наш взгляд, она имеет недостаток, который заключается в достаточно свободной интерпретации степени повреждения связочного аппарата. Для ее активного применения необходимо ясное понимание механизмов травмы и способность полно оценить повреждение связочного аппарата по косвенным признакам. Отсутствие четких критериев повреждения связочного комплекса, большая шкала оценки (от 1 до 20 баллов) и разный профессиональный уровень экспертов, казалось бы, должны были привести к большому разбросу ответов. Однако во всех работах были получены почти «идеальные» значения ИСС, который приближался к 1 при оценке согласованности заключений как одного и того же эксперта, так и разных экспертов. Значения были достаточно высокими не только у профессионалов, но и у резидентов. Каппу Флейса ни в одной работе не рассчитывали, поэтому оценить точность совпадения цифровых значений у разных наблюдателей не представляется возможным. Еще одним существенным ограничением исследований CSISS было участие экспертов одного и того же центра, что могло повлиять на результаты. Кроме того, ни в одной работе не был точно указан способ подсчета ИСС (с применением простых или средних мер).

Таким образом, проблема выбора шкалы/классификации повреждений НШОП остается актуальной. Согласно данным исследований, наибольшей надежностью и воспроизводимостью результатов на сегодняшний день характеризуются классификация AOSpine и шкала SLIC. Классификация Allen–Fergusson имеет более низкий коэффициент согласованности заключений экспертов, но при этом она способна дать более наглядное представление о повреждении. При условии

ее регулярного применения в условиях клиники каппа Флейса, отражающая согласованность заключений разных экспертов, могла быть существенно выше. Тем не менее никто не изучал кривую обучаемости и динамику показателей согласованности заключений на фоне процесса обучения применению данной шкалы. Так же поверхностно изучена классификация J. Harris и соавт. Надежность классификации C. Argenson и соавт. не исследована. Данных для полноценной оценки и интерпретации шкалы CSISS имеющиеся в литературе работы не представляют. Большинство исследований проведены в пределах 1 стационара и базируются на оценке специально отобранных снимков пациентов.

Анализ работ наглядно продемонстрировал необходимость более тщательного исследования всех шкал и классификаций. Подобное исследование должно быть проведено на базе нескольких стационаров с привлечением экспертов с разным уровнем опыта (от резидентов до опытных спинальных хирургов). Должна быть проанализирована не только воспроизводимость результатов применения отдельных классификаций/шкал, но и кривая обучаемости и взаимосвязь между шкалами/классификациями.

ОГРАНИЧЕНИЯ ИССЛЕДОВАНИЯ

Ограничениями нашей работы стали недостаточное количество исследований (даже самой распространенной классификации Allen–Fergusson посвящено всего 3 работы, классификации J. Harris и соавт. – 1), небольшой объем выборок данных КТ и МРТ в большинстве статей и неоднородность групп, часть из которых была отобрана экспертами, а другие были сформированы методом сплошной выборки. На каппу Флейса в определенной степени влияет распределение признаков, что может приводить к ее снижению, даже несмотря на высокий уровень согласованности заключений и точности индивидуальных оценок [19].

Ограничивает результаты нашего исследования и небольшой размер групп экспертов, а также их разнородность по уровню опыта. Оценка при помощи ИСС может быть недостоверной, если не учитывать размеры выборки и рейтинг экспертов.

ЗАКЛЮЧЕНИЕ

Проведенный анализ показал, что наибольшей надежностью и воспроизводимостью результатов характеризуются классификация AOSpine и шкала SLIC. Тем не менее имеющихся в литературе данных недостаточно для полноценного сравнения всех шкал и классификаций. Должны быть проведены многоцентровые работы по валидации классификаций НШОП, чтобы выбрать или разработать оптимальную для применения в ежедневной практике.

ЛИТЕРАТУРА / REFERENCES

- Walters B.C., Hadley M.N., Hurlbert R.J. et al. Guidelines for the management of acute cervical spine and spinal cord injuries: 2013 update. *Neurosurgery* 2013;60 Suppl 1:82–91. DOI: 10.1227/01.neu.0000430319.32247.7f.
- Chhabra H.S., Kaul R., Kanagaraju V. Do we have an ideal classification system for thoracolumbar and subaxial cervical spine injuries: what is the expert's perspective? *Spinal Cord* 2015;53(1):42–8. DOI: 10.1038/sc.2014.194.
- Audigé L., Bhandari M., Hanson B., Kellam J. A concept for the validation of fracture classifications. *J Orthop Trauma* 2005;19(6):404–6.
- Moher D., Liberati A., Tetzlaff J., Altman D.G., PRISMA Group. Preferred Reporting Items for Systematic reviews and Meta-Analyses: the PRISMA statement. *PLoS Med* 2009;6(7):e1000097. DOI: 10.1371/journal.pmed1000097.
- Landis J.R., Koch G.G. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159–74.
- Бурцев А.В. Структура субаксиальных повреждений шейного отдела позвоночника и соответствие им степени неврологических расстройств. *Бюллетень Восточно-Сибирского научного центра Сибирского отделения РАМН* 2011; 3(ч. 1):21–3. [Burtsev A.V. The structure of subaxial damages of the cervical spine and accordance to them of the degree of neurological disorders. *Bulleten Vostochno-Sibirskogo nauchnogo tsentra Sibirskogo otdeleniya RAMN = Bulletin of Eastern-Siberian Scientific Center* 2011;3(Pt 1):21–3. (In Russ.)].
- Бурцев А.В. Выбор оптимальной классификации и лечебного алгоритма при субаксиальных повреждениях шейного отдела позвоночника. *Бюллетень Восточно-Сибирского научного центра Сибирского отделения РАМН* 2012;4(ч. 2):21–5. [Burtsev A.V. Choice of optimal classification and treatment algorithm at subaxial injuries of the cervical segment of spine. *Bulleten Vostochno-Sibirskogo nauchnogo tsentra Sibirskogo otdeleniya RAMN = Bulletin of Eastern-Siberian Scientific Center* 2012;4(Pt 2):21–5. (In Russ.)].
- Губин А.В., Бурцев А.В. Классификации субаксиальных повреждений шейного отдела позвоночника. *Хирургия позвоночника* 2012;(2):8–15. [Gubin A.V., Burtsev A.V. Classification of subaxial cervical spine injuries. *Khirurgiya pozvonochnika = Spine Surgery* 2012;(2):8–15. (In Russ.)].
- Anderson P.A., Moore T.A., Davis K.W. et al.; Spinal Trauma Study Group. Cervical spine injury severity score. Assessment of reliability. *J Bone Joint Surg Am* 2007;89(5):1057–65. DOI: 10.2106/JBJS.F.00684.
- Van Middendorp J.J., Audigé L., Bartels R.H. et al. The Subaxial Cervical Spine Injury Classification System: an external agreement validation study. *Spine J* 2013;13(9):1055–63. DOI: 10.1016/j.spinee.2013.02.040.
- Silva O.T., Sabba M.F., Lira H.I. et al. Evaluation of the reliability and validity of the newer AOSpine subaxial cervical injury classification (C-3 to C-7). *J Neurosurg Spine* 2016;25(3):303–8. DOI: 10.3171/2016.2.SPINE151039.
- Stone A.T., Bransford R.J., Lee M.J. et al. Reliability of classification systems for subaxial cervical injuries. *Evid Based Spine Care J* 2010;1(3):19–26. DOI: 10.1055/s-0030-1267064.
- Urrutia J., Zamora T., Campos M. et al. A comparative agreement evaluation of two subaxial cervical spine injury classification systems: the AOSpine and the Allen and Ferguson schemes. *Eur Spine J* 2016;25(7): 2185–92. DOI: 10.1007/s00586-016-4498-0.
- Vaccaro A.R., Hulbert R.J., Patel A.A. et al.; Spine Trauma Study Group. The subaxial cervical spine injury classification system: a novel approach to recognize the importance of morphology, neurology, and integrity of the disco-ligamentous complex. *Spine (Phila Pa 1976)* 2007;32(21):2365–74. DOI: 10.1097/BRS.0b013e3181557b92.
- Vaccaro A.R., Koerner J.D., Radcliff K.E. et al. AOSpine subaxial cervical spine injury classification system. *Eur Spine J* 2016;25(7):2173–84. DOI: 10.1007/s00586-015-3831-3.
- Zehnder S.W., Lenarz C.J., Place H.M. Teachability and reliability of a new classification system for lower cervical spinal injuries. *Spine (Phila Pa 1976)* 2009;34(19):2039–43. DOI: 10.1097/BRS.0b013e3181af053c.
- Mirza S.K., Mirza A.J., Chapman J.R., Anderson P.A. Classifications of thoracic and lumbar fractures: rationale and supporting data. *J Am Acad Orthop Surg* 2002;10(5):364–77.
- Van Middendorp J.J., Audigé L., Hanson B. et al. What should an ideal spinal injury classification system consist of? A methodological review and conceptual proposal for future classifications. *Eur Spine J* 2010;19(8):1238–49. DOI: 10.1007/s00586-010-1415-9.
- Uebersax J.S. Validity inferences from interobserver agreement. *Psychol Bull* 1988;104(3):405–16.

Вклад авторов

А.А. Гринь: разработка дизайна исследования, анализ полученных данных, написание текста статьи;
И.С. Львов: анализ полученных данных, обзор публикаций по теме статьи, написание текста статьи;
С.Л. Аракелян: обзор публикаций по теме статьи, подготовка иллюстраций;
А.Э. Тальпов: разработка дизайна исследования, анализ полученных данных;
А.Ю. Кордонский: разработка дизайна исследования, анализ полученных данных, написание текста статьи;
А.В. Сытник: получение данных для анализа, обзор публикаций по теме статьи;
Б.А. Абдухаликов: получение данных для анализа, обзор публикаций по теме статьи;
У.Г. Хушназаров: получение данных для анализа, обзор публикаций по теме статьи;
В.А. Каранадзе: получение данных для анализа, обзор публикаций по теме статьи;
В.В. Крылов: разработка дизайна исследования.

Authors' contributions

A.A. Grin: developing the research design, analysis of the obtained data, article writing;
I.S. Lvov: analysis of the obtained data, reviewing of publications of the article's theme, article writing;
S.L. Arakelyan: reviewing of publications of the article's theme, creation of illustrations;
A.E. Talyov: developing the research design, analysis of the obtained data;
A.Yu. Kordonsky: developing the research design, analysis of the obtained data, article writing;
A.V. Sytnik: obtaining data for analysis, reviewing of publications of the article's theme;
B.A. Abdukhalikov: obtaining data for analysis, reviewing of publications of the article's theme;
U.G. Khushnazarov: obtaining data for analysis, reviewing of publications of the article's theme;
V.A. Karanadze: obtaining data for analysis, reviewing of publications of the article's theme;
V.V. Krylov: developing the research design.

ORCID авторов/ORCID of authors

А.А. Гринь/A.A. Grin: <https://orcid.org/0000-0003-3515-8329>

И.С. Львов/I.S. Lvov: <https://orcid.org/0000-0003-1718-0792>

С.Л. Аракелян/S.L. Arakelyan: <https://orcid.org/0000-0003-1381-2350>

А.Э. Талыпов/A.E. Talypov: <https://orcid.org/0000-0002-6789-8164>

А.Ю. Кордонский/A.Yu. Kordonsky: <https://orcid.org/0000-0001-5344-3970>

А.В. Сытник/A.V. Sytnik: <https://orcid.org/0000-0001-5565-4018>

Б.А. Абдухаликов/B.A. Abdukhalikov: <https://orcid.org/0000-0002-0131-8008>

У.Г. Хушназаров/U.G. Khushnazarov: <https://orcid.org/0000-0001-6981-3243>

В.А. Каранадзе/V.A. Karanadze: <https://orcid.org/0000-0003-0180-9154>

В.В. Крылов/V.V. Krylov: <https://orcid.org/0000-0001-5256-0905>

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest. The authors declare no conflict of interest.

Финансирование. Исследование проведено без спонсорской поддержки.

Financing. The study was performed without external funding.

Статья поступила: 17.10.2018. **Принята к публикации:** 30.01.2019.

Article received: 17.10.2018. **Accepted for publication:** 30.01.2019.