https://doi.org/10.47183/mes.2025-278

DEVELOPMENT OF PREDICTIVE MATHEMATICAL MODELS FOR PHYSICAL PERFORMANCE PARAMETERS IN SPORTS AND SPORTS MEDICINE

Victoria V. Petrova✉

Burnasyan Federal Medical Biophysical Center, Moscow, Russia

Introduction. Predictive modeling in healthcare is a rapidly evolving field of scientific knowledge at the intersection of information technology and medicine. In sports medicine, the importance of accurate forecasting of physical performance parameters in response to changing environmental conditions cannot be overstated. For athletes, such information provides a crucial competitive advantage before major competitions.

Objective. Development of methods and approaches to analyze clinical data obtained through comprehensive medical examinations of athletes.

Materials and methods. An analysis of anonymized medical data from comprehensive medical examinations was conducted for 6222 world-class athletes (3792 males and 2430 females) with a mean age of 23.3 ± 5.1 years. The data were stratified by sex and according to sports categories: cyclic sports (1376 athletes, including 861 males and 515 females); complex coordination sports (1342 athletes, including 761 males and 581 females); team sports (1618 athletes, including 980 males and 638 females); and combat sports (1886 athletes, including 1190 males and 696 females). The analysis included both clinical data on the presence (or absence) of pathological conditions identified during specialist medical examinations and physiological parameters from bicycle ergometer stress testing. Statistical analysis was performed using the StatTech v. 4.6.0 software (StatTech, Russia).

Results. Using regression analysis, statistically significant ($p < 0.001$) predictive models for a set of physical performance parameters were developed, which revealed over 40 associations with clinical diagnoses made by medical specialists. The strongest correlations were observed between physical performance indicators and dental diagnoses. Future research will focus on creating a mathematical model to predict performance decline in world-class athletes, based on an analysis of disease development risk factors.

Conclusions. The developed and implemented approaches for analyzing clinical data from comprehensive medical examinations of world-class athletes enabled the creation of effective predictive mathematical models of physical performance parameters using linear regression methodology, while accounting for the presence/absence of identified diagnoses. The proposed models provide a comprehensive assessment of athletes' functional status, thus allowing accurate prediction of physical performance levels and optimization of professional training by minimizing risks of overtraining and sports-related injuries.

Keywords: high-performance sports; mathematical model; physical performance parameters; pathological condition; regression analysis

For citation: Petrova V.V. Development of predictive mathematical models for physical performance parameters in sports and sports medicine. *Extreme Medicine*. 2025;27(3):392–399. https://doi.org/10.47183/mes.2025-278

Funding: the study was conducted without external funding.

Acknowledgements: the author expresses her gratitude to Prof. Alexander S. Samoilov, Member Correspondent of the Russian Academy of Sciences, for his assistance, advice, and valuable critical comments during work on this article.

Compliance with ethical principles: the study was approved by the local Ethics Committee of the Burnasyan Federal Medical Biophysical Center (protocol No. 121 dated 23.01.2025). All participants signed a voluntary consent form for the use of their anonymized medical data in scientific research.

Potential conflict of interest: the author declares no conflict of interest.

✉ Victoria V. Petrova vpetrova@fmbcfmba.ru

Received: 21 Feb. 2025 **Revised:** 3 June 2025 **Accepted:** 7 Aug. 2025 **Online first:** 20 Aug. 2025

УДК 61:796/799

РАЗРАБОТКА ПРОГНОСТИЧЕСКИХ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ПАРАМЕТРОВ ФИЗИЧЕСКОЙ РАБОТОСПОСОБНОСТИ В СПОРТЕ И СПОРТИВНОЙ МЕДИЦИНЕ

В.В. Петрова✉

Федеральный медицинский биофизический центр им. А.И. Бурназяна ФМБА России, Москва, Россия

Введение. Прогностическое моделирование в здравоохранении — новая развивающаяся отрасль научного знания, находящаяся на стыке информационных технологий и медицины. Для спортивной медицины наличие точного прогноза параметров физической работоспособности в ответ на изменяющиеся условия внешней среды сложно переоценить, а для спортсмена подобная информация даст необходимое конкурентное преимущество при проведении ответственных соревнований.

Цель. Разработка методов и подходов к анализу клинических данных углубленного медицинского обследования (УМО) спортсменов.

Материалы и методы. Проведен анализ обезличенных медицинских данных результатов УМО для 6222 спортсменов высокого класса (3792 мужчины и 2430 женщин) (средний возраст 23,3 ± 5,1 года). Данные были распределены по полу и в соответствии с группами видов спорта: циклические виды спорта (1376 спортсменов, из них 861 мужчина и 515 женщин); сложнокоординационные виды спорта (1342 спортсмена, из них 761 мужчина и 581 женщина); игровые виды спорта (1618 спортсменов, из них 980 мужчин и 638 женщин) и спортивные единоборства (1886 спортсменов, из них 1190 мужчин и 696 женщин). Анализу подверглись как клинические данные по наличию (отсутствию) нозологических единиц, выявленных в ходе осмотров врачами-специалистами, так и физиологические показатели нагрузочного тестирования на велоэргометре. Статистический анализ проведен с использованием программы StatTech v. 4.6.0 (разработчик — ООО «Статтех», Россия).

Результаты. В результате на основе метода регрессионного анализа были построены достоверные ($p < 0{,}001$) прогностические модели группы параметров физической работоспособности, которые выявили наличие более 40 связей с клиническими диагнозами врачей-специалистов. Больше всего взаимосвязей было зафиксировано между группой показателей физической работоспособности и проставленными диагнозами стоматолога. Дальнейшая работа будет направлена на разработку математической модели прогнозирования снижения результативности у спортсменов спорта высших достижений, основанной на анализе рисков развития заболеваний.

Выводы. Разработанные и примененные подходы к анализу клинических данных углубленного медицинского обследования спортсменов высокого класса позволили, применяя метод линейной регрессии, создать эффективные прогностические математические модели параметров физической работоспособности с учетом наличия/отсутствия выявленного диагноза. Предложенные модели обеспечивают комплексную оценку функционального состояния спортсменов, что способствует более точному прогнозированию уровня физической работоспособности и позволяет оптимизировать профессиональную деятельность, минимизируя риски перетренированности и травматизма.

Ключевые слова: спорт высших достижений; математическая модель; параметры физической работоспособности; нозологическая единица; регрессионный анализ

Для цитирования: Петрова В.В. Разработка прогностических математических моделей параметров физической работоспособности в спорте и спортивной медицине. *Медицина экстремальных ситуаций*. 2025;27(3):392–399. https://doi.org/10.47183/mes.2025-278

Финансирование: исследование выполнено без спонсорской поддержки.

Благодарность: Александру Сергеевичу Самойлову, д-ру мед. наук, профессору, чл.-корр. РАН, за помощь, консультацию и ценные критические замечания в процессе написания данной статьи.

Соответствие принципам этики: исследование одобрено локальным этическим комитетом ФГБУ «ГНЦ — ФМБЦ им. А.И. Бурназяна» ФМБА России (протокол № 121 от 23.01.2025). Всеми участниками подписано добровольное согласие на использование своих обезличенных медицинских данных в научных исследованиях.

Потенциальный конфликт интересов: автор заявляет об отсутствии конфликта интересов.

✉ Петрова Виктория Викторовна vpetrova@fmbcfmba.ru

Статья поступила: 21.02.2025 **После доработки:** 03.06.2025 **Принята к публикации:** 07.08.2025 **Online first:** 20.08.2025

INTRODUCTION

The application of mathematical models to predict and evaluate physiological parameters in professional athletes is closely related to both advancements in sports science and analytical methods.

In the field of sports performance analysis and forecasting, modern statistical modeling techniques are increasingly finding application, driving the transformation of research methodologies worldwide. While sports science has traditionally relied on conventional statistical approaches, recent innovations have introduced more sophisticated tools such as machine learning algorithms and hierarchical modeling. These advanced techniques enable researchers to identify complex relationships within both medical and athletic data, leading to deeper insights into the predictors of performance decline and the optimization of training strategies [1–6].

Previous research in this domain has primarily focused on understanding individual and collective trends in general and sport-specific physical performance metrics. For instance, studies on running performance have explored three key areas: (1) physiological determinants of world-record achievements; (2) development of equivalent scoring and race outcome prediction systems; and (3) modeling individual physiological parameters in track and field athletes [7–9]. Despite these efforts, there is a lack of comprehensive knowledge-based models capable of integrating and comparing all the above aspects. This determines the relevance of developing versatile, personalized, and accessible mathematical models for predicting physical performance parameters in sports.

Recent achievements in the integration and analysis of big data have enhanced the accuracy of performance predictions. In one study, the researchers used an online database of the performance of British athletes (1954–2013) to propose a simplified model capturing key performance characteristics while maintaining empirical validity. This model demonstrated a remarkably low mean prediction error for specific athletic outcomes, marking an important step toward unifying performance understanding through data analytics [10, 11].

A review of literature on predictive mathematical models in sports medicine reveals that most studies focus either on injury risk prediction across various sports or on estimating the probability of achieving target performance outcomes [12–15]. However, no studies examining potential correlations between diagnosed medical conditions in world-class athletes and their predicted impact on physical performance decline have been found.

With advances in neural networks and machine learning, researchers — including international collaborations — are now combining efforts to study the interplay of physiological and psychological factors affecting athletic performance. The goal is to optimize training regimens through predictive models capable of accounting for the dynamic nature of sports performance and its relationship with athletes' physiological processes.

The present study is aimed at developing methods for analyzing clinical data from comprehensive medical examinations (CME) of athletes.

MATERIALS AND METHODS

This study analyzed and mathematically processed anonymized clinical and instrumental examination data derived from medical records of world-class athletes who underwent comprehensive medical examinations (CME) in the Burnasyan Federal Medical Biophysical Center in 2019–2023. The dataset included clinical information from 6222 athletes (3792 males and 2430 females, mean age 23.3 ± 5.1 years).

The data were stratified by sex and sport categories: cyclic sports (1376 athletes, including 861 males and 515 females); complex coordination sports (1342 athletes, including 761 males and 581 females); team sports (1618 athletes, including 980 males and 638 females); and combat sports (1886 athletes, including 1190 males and 696 females).

The article presents data on male athletes from cyclic (n = 861) and team sports (n = 980). These cohorts were selected as those capable of providing the most representative and homogeneous samples for robust statistical analysis and reliable mathematical modeling. Data from other sports categories were utilized for preliminary screening and selection purposes.

During model development, the following inclusion criteria were applied:

- parameters had to reflect the key indicators of physical condition and functional capacity (morphofunctional status, aerobic/anaerobic capabilities, etc.);
- data were collected within homogeneous athlete groups (matched by sex, age, sports category, and skill level);
- measurements were obtained using standardized protocols (with uniform CME procedures for exercise testing).

The exclusion criteria comprised incomplete, erroneous, or anomalous parameters that could distort modeling results.

Thus, the following baseline parameters for mathematical modeling were used: age (years); weight (kg); height (cm); oxygen consumption at AT (anaerobic threshold) level ($VO_{2\ AT}$, mL/min/kg); peak oxygen consumption during exercise testing ($VO_{2\ peak}$, mL/min/kg); respiratory exchange ratio (R, relative units); resting heart rate (HR_{rest}, bpm); heart rate at aerobic threshold level (HR_{AerT}, bpm); heart rate at AT level (HR_{AT}, bpm); peak exercise heart rate (HR_{peak}, bpm); heart rate at 3 min of recovery (HR_{3min}, bpm); power output at anaerobic threshold level (Pwr_{AT}, W); peak power output during testing (Pwr_{peak}, W); relative power output at anaerobic threshold level (Pwr_{AT}/weight, W/kg); relative peak power output during testing (Pwr_{peak}/weight, W/kg). In each sports category, the statistical dataset of the studied parameters was combinatorially grouped by qualitative presence (1)/absence (0) of pathological conditions diagnosed by the following medical specialists: gastroenterologist, dermatovenerologist, cardiologist, neurologist, otolaryngologist, ophthalmologist, dentist, orthopedic traumatologist, and endocrinologist.

All quantitative parameters in the modeling are presented as: X_M — sex (0 — female, 1 — male); X_{age} — age; X_{weight} — weight; X_{height} — height; X_{gastro} — pathological condition by gastroenterologist (0 — no, 1 — yes); X_{derm} — pathological condition by dermatovenerologist (0 — no, 1 — yes); X_{cardio} — pathological condition by cardiologist (0 — no, 1 — yes); X_{neuro} — pathological condition by neurologist (0 — no, 1 — yes); X_{oto} — pathological condition by otolaryngologist (0 — no, 1 — yes); X_{ophth} — pathological condition by ophthalmologist (0 — no, 1 — yes); X_{dent} — pathological condition by dentist (0 — no, 1 — yes); X_{trauma} — pathological condition by orthopedic traumatologist (0 — no, 1 — yes); X_{endo} — pathological condition by endocrinologist (0 — no, 1 — yes); $X_{VO2\ AT}$ — $V(O_2)$ at AT level; $X_{VO2\ peak}$ — $V(O_2)$ at peak; X_R — respiratory coefficient; $X_{HR\ rest}$ — HR at rest; $X_{HR\ AerT}$ — HR at AerT; $X_{HR\ AT}$ — HR at AT; $X_{HR\ peak}$ — HR at peak; $X_{HR\ 3min}$ — HR at 3-min recovery; $X_{PWR\ AT}$ — power at AT; $X_{PWR\ peak}$ — power at peak; $X_{PWR\ AT/weight}$ — power at AT divided by weight; $X_{PWR\ peak/weight}$ — power at peak divided by weight.

Statistical analysis was performed using the StatTech v. 4.6.0 software (StatTech, Russia). The linear regression method was used to examine the dataset structure and establish relationships between its parameters. Mathematical models were developed to describe dependencies between the group of quantitative physical performance indicators and the presence/absence of pathological conditions identified by physicians during in-depth medical examinations, for each sport category and sex. Regression equation coefficients were determined using the least squares method with Cramer's formulas.

RESULTS AND DISCUSSION

During data preparation for linear regression modeling, the relevant CME results of athletes were compiled to ensure the required precision. Incomplete, erroneous, or anomalous values were excluded, along with noninformative features. Table 1 presents the general characteristics of male athletes from cyclic and team sports included in the final sample.

Among the exercise testing parameters characterizing the overall physical performance (items 4–15 in Table 1), the most physiologically relevant indicators for sports medicine applications are those of gas exchange: $VO_{2\ AT}$, $VO_{2\ peak}$, respiratory exchange ratio, as well as the direct measure of achieved power output at anaerobic threshold.

The models presented below describe general relationships (and interdependencies) between the selected physical performance metrics (Y value in the formula) and all other parameters, including the presence/absence of pathological conditions during examination (X values in the formula).

The observed relationship for oxygen consumption at anaerobic threshold $VO_{2\ AT}$ (1 — cyclic sports, 2 — team sports) is described by the following linear regression equations:

$$Y_{VO2\ AT} = -5.313 - 0.424 \times X_{neuro} - 0.369 \times X_{dent} + 0.921 \times X_{endo} + 0.058 \times X_{weight} + 0.479 \times X_{VO2\ peak} - 0.024 \times X_{HR\ rest} + 0.029 \times X_{HR\ AT} + 0.038 \times X_{PWR\ AT} - 0.047 \times X_{PWR\ peak} + 5.924 \times X_{PWR\ AT/weight}, \quad (1)$$

$$Y_{VO2\ AT} = -0.301 + 0.572 \times X_M + 0.576 \times X_{VO2\ peak} - 0.022 \times X_{HR\ rest} + 0.072 \times X_{HR\ AT} - 0.033 \times X_{HR\ peak} - 0.012 \times X_{HR\ 3min} - 0.004 \times X_{PWR\ peak} + 7.671 \times X_{PWR\ AT/weight} - 3.725 \times X_{PWR\ peak/weight}. \quad (2)$$

The obtained regression models for oxygen consumption at anaerobic threshold ($VO_{2\ AT}$) demonstrated the following characteristics:

1. The multiple correlation coefficient was $R_{xy} = 0.965$ for cyclic sports and $R_{xy} = 0.948$ for team sports, indicating a highly strong relationship according to the Chaddock scale;

2. The coefficient of multiple determination was $R^2 \approx (0.965)^2 = 0.931$ for cyclic sports and $R^2 \approx (0.948)^2 = 0.899$ for team sports. These models can predict $VO_{2\ AT}$ values with high accuracy: they explain 93.1% of observed variance in cyclic sports and 89.9% in team sports. The models were statistically significant ($p < 0.001$).

After accounting for interdependencies among physical performance parameters, negative associations were found between $VO_{2\ AT}$ and neurological/dental pathological conditions, while a positive association was observed with endocrine disorders.

The relationship for peak oxygen consumption ($VO_{2\ peak}$) (3 — cyclic sports, 4 — team sports) is described by the following linear regression equation:

$$Y_{VO2\ peak} = 5.920 + 0.737 \times X_M - 1.417 \times X_{endo} + 0.865 \times X_{VO2\ AT} - 0.050 \times X_{HR\ AT} + 0.036 \times X_{HR\ peak} - 6.026 \times X_{PWR\ AT/weight} + 6.668 \times X_{PWR\ peak/weight}, \quad (3)$$

$$Y_{VO2\ peak} = 5.743 + 5.743 \times X_M - 0.091 \times X_{age} + 0.823 \times X_{VO2\ AT} - 6.022 \times X_R + + 0.023 \times X_{HR\ rest} - 0.076 \times X_{HR\ AT} + 0.048 \times X_{HR\ peak} - 6.090 \times X_{PWR\ AT/weight} + + 8.022 \times X_{PWR\ peak/weight}. \quad (4)$$

The key characteristics of the developed models:

1. The multiple correlation coefficient was $R_{xy} = 0.933$ for cyclic sports and $R_{xy} = 0.919$ for team sports, indicating a highly strong association according to the Chaddock scale.

2. The coefficient of determination reached $R^2 \approx (0.933)^2 = 0.871$ for cyclic sports and $R^2 \approx (0.919)^2 = 0.844$ for team sports.

The models demonstrate high predictive accuracy for $VO_{2\ peak}$ values, explaining 87.1% of the observed variance in cyclic sports and 84.4% in team sports. All models showed statistical significance ($p < 0.001$).

After analyzing interdependencies among physical performance parameters, a negative association between $VO_{2\ peak}$ and endocrine disorders was observed.

The relationship between the respiratory exchange ratio (R) (5 — cyclic sports, 6 — team sports) and quantitative factors is described by the following linear regression equation:

$$Y_R = 1.436 - 0.013 \times X_M + 0.013 \times X_{dent} - 0.002 \times X_{height} - 0.002 \times X_{weight} - 0.0001 \times X_{HR\ rest} + 0.0001 \times X_{HR\ AerT} - 0.001 \times X_{HR\ AT} + 0.001 \times X_{HR\ peak} + 0.001 \times X_{HR\ 3min} + 0.001 \times X_{PWR\ AT} - 0.167 \times X_{PWR\ AT/weight} + 0.107 \times X_{PWR\ peak/weight}, \quad (5)$$

$$Y_R = 0.388 + 0.388 \times X_M + 0.012 \times X_{ophth} + 0.009 \times X_{dent} + 0.002 \times X_{weight} - 0.002 \times X_{VO2peak} - 0.001 \times X_{HR\ AT} + + 0.000 \times X_{HR\ 3min} + 0.001 \times X_{HR\ peak} - 0.001 \times X_{PWR\ AT} + 0.092 \times X_{PWR\ peak/weight}. \quad (6)$$

The characteristics of the developed models are as follows:

1. For cyclic sports, the multiple correlation coefficient was $R_{xy} = 0.830$, while for team sports it was $R_{xy} = 0.783$, indicating a strong association according to the Chaddock scale.

2. The coefficient of determination was $R^2 \approx (0.830)^2 = 0.689$ for cyclic sports and $R^2 \approx (0.783)^2 = 0.613$ for team sports.

The models demonstrate moderately high predictive accuracy for the respiratory exchange ratio (R), explaining 68.9% of the observed variance in cyclic sports and

Table 1. Descriptive statistics of quantitative variables included in the analysis

No.	Examined parameter	Median (mean) parameter value	Value range	
			min	max
Cyclic sports (*n* = 861)				
1	Age, years	21.00 [19.0–25.0]	18.00	26.00
2	Height, cm	180.00 [172.0–186.0]	152.00	207.00
3	Weight, kg	74.00 [65.0–83.0]	43.00	120.00
4	$VO_{2\ AT}$, mL/min/kg	42.68 ± 9.71 (42.16–43.19)	14.74	68.36
5	$VO_{2\ peak}$, mL/min/kg	49.97 [43.99–57.25]	2.52	92.93
6	*R*, relative units	1.16 [1.10–1.23]	0.89	1.55
7	HR_{rest}, bpm	76.00 [67.00–85.00]	40.00	126.00
8	HR_{AerT}, bpm	120.00 [108.00–132.00]	61.00	175.00
9	HR_{AT}, bpm	155.00 [144.00–165.00]	95.00	196.00
10	HR_{peak}, bpm	173.00 [164.00–181.00]	18.00	206.00
11	HR_{3min}, bpm	102.00 [92.00–114.00]	45.00	155.00
12	Pwr_{AT}, W	275.00 [225.00–340.00]	90.00	520.00
13	Pwr_{peak}, W	345.00 [285.00–420.00]	115.00	600.00
14	Pwr_{AT}/weight, W/kg	3.81 ± 0.90 (3.76–3.85)	0.00	6.34
15	Pwr_{peak}/weight, W/kg	4.69 [4.15–5.43]	1.42	7.72
Team sports (*n* = 980)				
1	Age, years	22.00 [19.00–26.00]	18.00	26.00
2	Height, cm	182.00 [173.00–191.00]	151.00	220.00
3	Weight, kg	80.00 [68.00–92.00]	47.00	126.00
4	$VO_{2\ AT}$, mL/min/kg	33.42 [28.92–38.27]	14.26	58.86
5	$VO_{2\ peak}$, mL/min/kg	41.25 ± 6.84 (40.92–41.58)	15.85	69.41
6	*R*, relative units	1.14 [1.09–1.19]	0.90	1.52
7	HR_{rest}, bpm	79.00 [71.00–86.00]	44.00	142.00
8	HR_{AerT}, bpm	117.00 [107.00–128.00]	67.00	177.00
9	HR_{AT}, bpm	149.00 [137.00–159.00]	91.00	199.00
10	HR_{peak}, bpm	168.00 [159.00–176.00]	65.00	202.00
11	HR_{3min}, bpm	103.00 [93.00–112.00]	29.00	173.00
12	Pwr_{AT}, W	235.00 [195.00–285.00]	80.00	470.00
13	Pwr_{peak}, W	310.00 [245.00–365.00]	130.00	525.00
14	Pwr_{AT}/weight, W/kg	2.99 [2.57–3.44]	0.00	5.27
15	Pwr_{peak}/weight, W/kg	3.85 ± 0.63 (3.82–3.89)	1.91	6.18

Table prepared by the author using her own data

Note: $VO_{2\ AT}$, $VO_{2\ peak}$, Pwr_{AT}/weight, Pwr_{Peak}/weight are presented as mean ± standard error of the mean (*M* ± SEM); all other parameters are presented as median (M_e) with lower and upper quartiles *Q* [25–75%]; $VO_{2\ AT}$ — oxygen consumption at anaerobic threshold level, $VO_{2\ peak}$ — oxygen consumption at maximal exercise testing stage, *R* — respiratory exchange ratio, HR_{rest} — heart rate at rest (pre-exercise), HR_{AerT} — heart rate at aerobic threshold level, HR_{AT} — heart rate at anaerobic threshold level, HR_{peak} — peak heart rate during exercise, HR_{3min} — heart rate at 3 min of recovery, Pwr_{AT} — power output at anaerobic threshold level, Pwr_{Peak} — relative maximal power output during testing, Pwr_{AT}/weight — relative power output at anaerobic threshold per body weight, Pwr_{Peak}/weight — relative maximal power output per body weight during testing.

61.3% in team sports. All models were statistically significant ($p < 0.001$).

After controlling for interdependencies among physical performance parameters, positive associations were identified between the respiratory exchange ratio (R) and ophthalmological/dental pathological conditions.

The relationship between power output at anaerobic threshold (Pwr_{AT}) (7 — cyclic sports, 8 — team sports) and quantitative factors is described by the following linear regression equation:

$$Y_{\text{PWR AT}} = -62.319 + 0.115 \times X_{\text{height}} + 0.274 \times X_{\text{weight}} + 0.167 \times X_{VO2\ \text{AT}} + 17.519 \times X_R - 0.092 \times X_{\text{HR AerT}} + \\ + 0.153 \times X_{\text{HR AT}} - 0.069 \times X_{\text{HR peak}} + 0.724 \times X_{\text{PWR peak}} + 74.556 \times X_{\text{PWR AT/weight}} - 55.694 \times X_{\text{PWR peak/weight}}, \quad (7)$$

$$Y_{\text{PWR AT}} = -49.907 + 1.485 \times X_{\text{neuro}} + 1.185 \times X_{\text{офт}} - 0.102 \times X_{\text{age}} + 0.089 \times X_{\text{height}} + + 0.434 \times X_{\text{weight}} + \\ + 0.154 \times X_{\text{HR AT}} - 0.159 \times X_{\text{HR peak}} + 76.528 \times X_{\text{PWR AT/weight}} + 0.652 \times X_{\text{PWR peak}} - 48.533 \times X_{\text{PWR peak/weight}}. \quad (8)$$

The characteristics of the developed models are as follows:

1. The multiple correlation coefficient was $R_{xy} = 0.996$ for cyclic sports and $R_{xy} = 0.994$ for team sports, indicating an exceptionally strong association according to the Chaddock scale.
2. The coefficient of determination was $R^2 \approx (0.996)^2 = 0.993$ for cyclic sports and $R^2 \approx (0.994)^2 = 0.988$ for team sports. The models demonstrate extremely high predictive accuracy for power output at anaerobic threshold (Pwr_{AT}), explaining 99.3% of the observed variance in cyclic sports and 98.8% in team sports. All models were statistically significant ($p < 0.001$).

After controlling for interdependencies among physical performance parameters, positive associations were identified between Pwr_{AT} and neurological/ophthalmological pathological conditions.

In a similar manner, clinical data from 6222 athletes (accounting for gender and sport category differences) were processed. After developing 12 separate models for each sports category, significant relationships (both

Table 2. Statistically significant associations between physical performance parameters and clinically diagnosed medical conditions

Parameter / Medical specialist	$VO_{2\ AT}$	$VO_{2\ peak}$	R	HR_{rest}	HR_{AerT}	HR_{AT}	HR_{peak}	HR_{3min}	Pwr_{AT}	Pwr_{peak}	Pwr_{AT}/weight	Pwr_{peak}/weight
Gastroenterologist			1	1	1				1	1		1
Dermatovenerologist	1	1		1					1			
Cardiologist			2		1							
Neurologist	1								1		1	2
Otolaryngologist				1				1				
Ophthalmologist			2	1	2				1		1	1
Dentist	1		2	1	1	1		1	1	1	1	1
Orthopedic traumatologist					1			1				
Endocrinologist	2	2			1							

Table prepared by the author using her own data

Note: "1" — relationship between the relevant physical performance indicators and the presence of specialist-diagnosed conditions in one sport category; "2" — relationship between the relevant physical performance indicators and the presence of specialist-diagnosed conditions in two sports categories simultaneously; $VO_{2\ AT}$ — oxygen consumption at anaerobic threshold level, $VO_{2\ peak}$ — oxygen consumption at maximal exercise testing stage, R — respiratory exchange ratio, HR_{rest} — heart rate at rest (pre-exercise), HR_{AerT} — heart rate at aerobic threshold level, HR_{AT} — heart rate at anaerobic threshold level, HR_{peak} — heart rate at peak exercise, HR_{3min} — heart rate at 3 minutes of recovery, Pwr_{AT} — power output at anaerobic threshold level, Pwr_{Peak} — relative maximal power output during testing, Pwr_{AT}/weight — relative power output at anaerobic threshold normalized to body weight, Pwr_{Peak}/weight — relative maximal power output normalized to body weight during testing.

positive and negative) between the studied parameters and the presence/absence of diagnosed pathological conditions were quantified.

The total number of significant associations between physical performance parameters and the presence/absence of documented pathological conditions across all studied athlete groups reached 46 (Table 2).

Out of 46 established dependencies, the strongest correlations were found with ophthalmologist-diagnosed conditions (8 confirmed relationships) and dentist-diagnosed conditions (11 confirmed relationships). The developed mathematical models for predicting physical performance parameters confirmed a strong relationship between the probability of achieving planned training results or performance in major competitions and the presence of specialist-diagnosed conditions identified during the comprehensive medical examinations of athletes.

CONCLUSION

In this study, effective predictive mathematical models of physical performance parameters, accounting for the presence/absence of diagnosed conditions, have been developed using clinical data from comprehensive medical examinations of world-class athletes and linear regression methods. The developed models can be used to carry out a comprehensive assessment of athletes' functional status, thus facilitating a more accurate prediction of physical performance levels and optimization of professional activities while minimizing risks of overtraining and injuries.

Thus, the study results contribute to the development of sports medicine and provide a scientific basis for decision making in the field of athlete preparation and medical support. This will subsequently lead to changes in sports physicians' approaches to interpreting the results of comprehensive medical examinations.

References

1. Badtieva VA, Sharykin AS, Pavlov VI. *Sports cardiology. Guide for cardiologists, pediatricians, doctors of functional diagnostics and sports medicine, trainers.* Moscow: IKAR, 2017 (In Russ.).
2. Mazitova GI. The role of the study of hemodynamic characteristics of rest in the prognosis of physical performance. *Theory and Practice of Physical Education.* 2008;1:83–5 (In Russ.).
EDN: NBMIJL
3. Petrova VV, Kish AA, Bragin MA. Forecast of physical performance based on indicators of a comprehensive assessment of the athletes' condition. *Medical Science and Education of the Urals*. 2019;1(97):155–9 (In Russ.).
EDN: ZDRHUT
4. Petrova VV, Nazarjan SE, Kish AA, Orlova NZ, Prudnikov IA. Analysis of existing methods for assessing the psychological state of an athlete to identify the most informative indicators that affect his performance. *Bulletin of Neurology, Psychiatry and Neurosurgery*. 2017;9:43–56 (In Russ.).
5. Rojas-Valverde D, Gómez-Ruano MA, Ibañez SJ, Nikolaidis PT. Editorial: New technologies and statistical models applied to sports and exercise science research: methodological, technical and practical considerations. *Frontiers in Sports and Active Living.* 2023;5:1267548.
https://doi.org/10.3389/fspor.2023.1267548
6. Wang Z, Zhang Q, Lan K, Yang Z, Gao X, Wu A, et al. Enhancing instantaneous oxygen uptake estimation by non-linear model using cardio-pulmonary physiological and motion signals. *Frontiers in Physiology*. 2022;13:1–14.
https://doi.org/10.3389/fphys.2022.897412
7. Bakhmayer M, Smolensky AV, Mityushkina OA. Professional risks in high-performance sports. *Bulletin of New Medical Technologies. Electronic edition.* 2015;3:9 (In Russ.).
EDN: UJXHPZ
8. Praskurnichij EA, Justova VD. Methodology for assessing cardiovascular risk in civil aviation flight personnel. *Aerospace and Environmental Medicine.* 2020;54(5):50–6 (In Russ.).
EDN: CHLUFU
9. Heazlewood T. Prediction versus reality: the use of mathematical models topredict elite performance in swimming and athletics at the Olympic games. *Journal of Sports Science and Medicine.* 2006;5(4):480–7.
10. Blythe DAJ, Király FJ. Prediction and Quantification of Individual Athletic Performance of Runners. *PLOS ONE.* 2016;11(6):e0157257.
https://doi.org/10.1371/journal.pone.0157257
11. Brown AW, Kaiser KA, Allison DB. Issues with data and analyses: errors, underlying themes, and potential solutions. *Proceedings of the National Academy of Science of the USA.* 2018;115(11):2563–70.
https://doi.org/10.1073/pnas.1708279115
12. Le Hyaric A, Aftalion A, Hanley B. Modelling the optimization of world-class 400 m and 1,500 m running performances using high-resolution data. *Frontiers in Sports and Active Living.* 2024;6:1293145.
https://doi.org/10.3389/fspor.2024.1293145
13. Fudin NA, Hadarcev AA. Medical and biological support for physical education and high-performance sports. *Bulletin of New Medical Technologies.* 2010;1:149–50 (In Russ.).
EDN: NWGJLN
14. Carey DL, Ong K, Whiteley R, Crossley KM, Crow J, Morris ME. Predictive modelling of training loads and injury in Australian football. *International Journal of Computer Science in Sport.* 2018;17(1):49–66.
https://doi.org/10.2478/ijcss-2018-0002
15. Khan NJ, Ahamad G, Naseem M, Sohail SS. Computational efficiency in sports talent identification — systematic review. *International Journal of Applied Decision Sciences.* 2023;16(3):358–84.
https://doi.org/10.1504/IJADS.2023.130600

Author contribution. The author confirms compliance with the ICMJE authorship criteria. Victoria V. Petrova — development of the work concept, writing the article text, critical review of the article content, making necessary corrections, and approval of the final version of the article for publication.

AUTHOR

Victoria V. Petrova, Cand. Sci. (Med.)
https://orcid.org/0000-0002-9987-6816
vpetrova@fmbcfmba.ru