



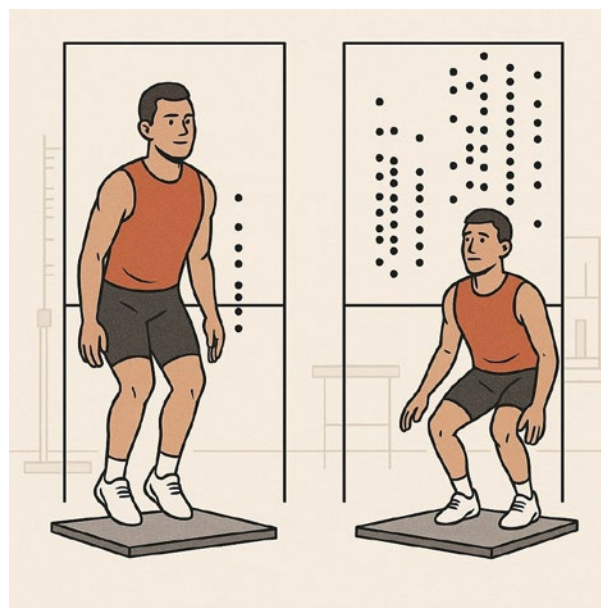
Žiga Kozinc¹

Testiranje v športni praksi: prepoznavanje resničnega napredka na individualni ravni

Izvleček

V članku obravnavamo pomen ponovljivosti pri testiranju športnikov in predstavimo dva ključna koraka za zanesljivo presojo napredka: najprej izbiro ustreznih testov na podlagi relativne in absolutne ponovljivosti iz literature, nato pa individualno oceno sprememb pri posamezniku. Relativna ponovljivost, najpogostejše izražena z intraklasnim koeficientom korelacije (angl. intraclass correlation coefficient – ICC), kaže, kako velike so napake pri ponovljenem testiranju glede na celotno variabilnost v skupini športnikov. Absolutna ponovljivost, ki se izraža med drugim s standardno napako merjenja (angl. standard error of measurement – SEM), tipično napako (angl. typical error – TE) in koeficientom variacije (angl. coefficient of variation – CV), pa navaja tipično vrednost napak, ki jo lahko v povprečju pričakujemo pri ponovitvi testa. V drugem delu članka opisujemo tri metode za ocenjevanje ponovljivosti meritev in oceno napredka na individualni ravni: (1) statistični model za ugotavljanje, ali sprememba statistično značilno presega kritično razliko; (2) metodo CV, ki oceni, ali sprememba presega običajno variabilnost posameznika; in (3) najmanjšo vredno spremembo (angl. smallest worthwhile change – SWC) za oceno praktične pomembnosti. Članek ponuja praktično orodje za trenerje pri odločanju, ali je napredek športnika resničen in pomemben.

Ključne besede: testiranje športnikov, ponovljivost meritev, relativna ponovljivost, absolutna ponovljivost.



Performance Testing in Sports: Assessing Meaningful Individual Progress

Abstract

The article discusses the importance of reliability in athlete testing and outlines two key steps for a trustworthy assessment of progress: first, the selection of appropriate tests based on relative and absolute reliability reported in the literature; and second, the individual evaluation of changes in each athlete. Relative reliability, most commonly expressed by the intraclass correlation coefficient (ICC), indicates the magnitude of measurement error in relation to the overall variability within a group of athletes. Absolute reliability, expressed through metrics such as the standard error of measurement (SEM), typical error (TE), and coefficient of variation (CV), reflects the typical amount of error expected when repeating a test. In the second part of the article, we present three methods for assessing measurement reliability and progress at the individual level: (1) a statistical model to determine whether a change statistically exceeds the critical difference; (2) the CV method, which evaluates whether a change surpasses the typical variability of the individual; and (3) the smallest worthwhile change (SWC) approach for judging practical significance. The article provides a practical tool for coaches to decide whether an athlete's progress is real and meaningful.

Keywords: Athlete testing, measurement reliability, relative reliability, absolute reliability.

¹ Fakulteta za vede o zdravju, Univerza na Primorskem, Izola, Slovenija

■ Uvod

Trenerji in športni strokovnjaki se pogosto srečujejo v vprašanju: »Ali je moj športnik napredoval?« Odgovor nanj običajno vsaj v določeni meri iščejo s ponavljanjem testov zmogljivosti, denimo s testiranjem največje porabe kisika, višine in daljine skokov, časov sprintov, meritev jakosti in moči ter drugih sposobnosti, pomembnih za posamezen šport. Rezultate pogosto primerjajo med športniki znotraj ekipe ali pa z referenčnimi vrednostmi, bodisi iz literature bodisi iz lastne zbirke podatkov. V zadnjih letih se je v športni znanosti in praksi okrepila zavest o pomembnosti individualnega pristopa (Jiménez-Reyes idr., 2017), ne le v treniranju, temveč tudi pri analizi podatkov (Harry idr., 2024). Ustrezno izbrani statistični pristopi nam omogočajo, da v raziskovanju bolj objektivno presodimo, ali je sprememba pri posamezniku resnična in vredna upoštevanja, ali pa gre le za naravno variacijo, ki se pojavi pri vsakem testiranju. V športni znanosti se za oceno napredka pogosto opiramo na preproste statistične metode, kot so t-testi in analiza variance (ANOVA). Te metode temeljijo na predpostavkah, kot so zadostno število preiskovancev, normalna porazdelitev podatkov in homogenost varianc. Vendar pa te pogoje v praksi le redko izpolnimo, posebno v okoljih z manjšimi ekipami ali v individualnih športih. Še pomembnejša omejitev pa je, da te metode temeljijo na povprečjih, nas pa v praksi v nasprotju z raziskavami najpogosteje zanima, kaj se dogaja z vsakim posameznikom posebej. Trener potrebuje orodje, ki mu pomaga ugotoviti, ali je sprememba v rezultatu testa pri posameznem športniku resnična, in ne le posledica merilnega šuma ali vsakodnevne variabilnosti (Bates, 1996).

Pri izbiri testov pogosto posegamo po tistih, ki so v literaturi označeni kot zanesljivi oziroma ponovljivi. Vendar je treba opozoriti, da izračuni ponovljivosti, ki jih najdemo v literaturi, izhajajo iz podatkov večjega vzorca in ne zagotavljajo, da bo test enako ponovljiv tudi pri posameznem športniku, s katerim delamo. Trener bi moral znati presoditi, ali je izbrani test dovolj stabilen tudi za tega konkretnega športnika in ali je opažena sprememba v rezultatu dovolj velika, da

jo lahko štejemo za resnični napredek. Brez tega tvega, da napačno oceni stanje športnika: morda zazna »napredek«, ki ga v resnici ni bilo, ali pa spregleda dejansko izboljšavo, ki bi jo bilo vredno upoštevati pri nadaljnjem treningu. V tem članku bomo v prvem delu povzeli najpogostejše mere ponovljivosti, ki lahko trenerjem pomagajo pri izbiri testov na podlagi znanstvene in strokovne literature, v drugem delu pa predstavili metode analize podatkov, s katerimi lahko v praksi presodimo o ponovljivosti testiranja za posameznega športnika in posledično o pomembnosti morebitnega napredka.

■ Izbira najprimernejših testov na podlagi literature

Relativna ponovljivost

Ko ocenjujemo, ali je posamezni test dovolj ponovljiv za uporabo v praksi, moramo upoštevati dve ključni dimenziji ponovljivosti: relativno in absolutno ponovljivost (Hopkins, 2000). Relativna ponovljivost opisuje, kako dosledno test razvršča posameznike znotraj skupine, torej ali posamezniki ohranjajo svojo relativno pozicijo (rang) znotraj skupine med več ponovitvami testiranja. Če športnik v prvem testiranju doseže najboljši rezultat, nas zanima, ali bo tudi v naslednjih ponovitvah ostal med najboljšimi. Relativno ponovljivost najpogosteje ocenjujemo z intraklasnim koeficientom korelacije (angl. intra-class correlation coefficient – ICC), ki izraža razmerje med variabilnostjo med posamezniki in celotno variabilnostjo podatkov. Višje vrednosti ICC pomenijo boljše relativno ponovljivost, torej večjo stabilnost rangiranja posameznikov znotraj skupine, in majhno variabilnost med meritvami posameznika v odnosu do skupne variabilnosti podatkov.

Po smernicah (Koo in Li, 2016) se vrednosti ICC interpretirajo tako, da vrednosti pod 0,50 kažejo na slabo ponovljivost, vrednosti med 0,50 in 0,75 na zmerno, vrednosti med 0,75 in 0,90 na dobro, vrednosti nad 0,90 pa na odlično ponovljivost. Močno priporočamo, da se poleg srednje vrednosti ICC vedno upošteva tudi 95-odstotni interval zaupanja. Interval zaupanja nam pove, znotraj katerega

razpona je z visoko verjetnostjo (95 %) dejanska (in ne le ocenjena) vrednost ICC. Če je ta razpon širok, je negotovost ocene velika, kar pomeni, da je lahko test v resnici precej slabše ponovljiv, kot kaže srednja vrednost ICC. Posebej previdni moramo biti, kadar spodnja meja intervala pade pod prag zmerne ponovljivosti (npr. pod 0,50), saj v takem primeru ne moremo z gotovostjo trditi, da gre za ponovljivi test. Nasprotno pa ozki intervali, ki so v celoti nad 0,75, dajejo večjo statistično gotovost in praktično zaupanje v ponovljivost testa. Prav tako je treba poudariti, da je vrednost ICC močno odvisna tudi od razpršenosti rezultatov v celotnem vzorcu. Če bodo posamezniki med seboj zelo različni (na primer pri raziskavi, v katero smo vključili osebe z različno stopnjo zmogljivosti), bodo vrednosti ICC kljub višjim napakam na ravni posameznika lahko nakazovale dobro ponovljivost. In obrnjeno, v zelo homogeni skupini bo ICC lahko nizek, čeprav so razlike med zaporednimi ponovitvami zelo majhne (na primer v skupini elitnih sprinterjev, ki imajo zelo podobne čase, rezultat posameznika pa lahko nekoliko variira). Zato je ključno, da v literaturi poleg vrednosti ICC za posamezen test poiščemo še kazalnike absolutne ponovljivosti. Poleg ICC občasno (v športni znanosti sicer redko) zasledimo še koeficient skladnosti (angl. concordance correlation coefficient – CCC), ki ni odvisen od variabilnosti med osebami, sicer pa se interpretira enako kot ICC (King idr., 2007). Korelacijski koeficienti, kot sta Pearsonov in Spearmanov koeficient, za oceno ponovljivosti niso ustrezni (Bunting idr., 2019).

Absolutna ponovljivost

Absolutna ponovljivost se nanaša na povprečno velikost napake pri ponovitvah testa pri enem samem posamezniku. Pove nam, kolikšen je tipičen razpon nihanja rezultatov, ki jih lahko pričakujemo pri istem športniku, če test večkrat ponovi. Z drugimi besedami, gre za oceno merilnega šuma oziroma naključnih odstopanj, ki niso posledica dejanske spremembe v zmogljivosti (Hopkins, 2000). Absolutna ponovljivost je ključna pri odločanju, ali je sprememba v rezultatu dovolj velika za sklepanje, da ni le posledica merilne napake. Absolutno ponovljivost najpogosteje izražamo s

standardno napako merjenja (angl. standard error of measurement – SEM), ki je izražena v enotah testa (npr. s cm, N in tako naprej). Pogosto se uporablja tudi tipična napaka (angl. typical error – TE). Vrednosti SEM zajamejo sistematično napako (glej razlago v nadaljevanju) in naključno napako, TE pa odraža samo naključno napako (merilni šum). Tako SEM kot TE se za lažjo interpretacijo lahko izrazi kot odstotek povprečne vrednosti, s čimer pridobimo koeficient variacije (angl. coefficient of variation – CV). Najpogosteje uporabljeni prag sprejemljive absolutne ponovljivosti je $CV < 10\%$ (Hopkins, 2000), najdemo pa tudi drugačna priporočila. Na področju športa so predlagali, da se $CV < 5\%$ šteje za dobro ponovljivost, $CV 5\text{--}10\%$ za sprejemljivo ponovljivost, $CV < 15\%$ pa za nesprejemljivo ponovljivost (Banyard idr., 2018).

Sistematična napaka

Poleg ocene relativne in absolutne variabilnosti med ponovitvami testiranja je pri presoji ponovljivosti pomembno upoštevati tudi možnost sistematične napake (Barraza idr., 2019). Gre za pojav, pri katerem se rezultati v ponovitvah ne razlikujejo le naključno, temveč kažejo enosmerno spremembo; na primer, če so rezultati drugega testiranja dosledno višji (v športu pogosto zaradi učinka učenja) ali pa dosledno nižji (na primer zaradi pojava utrujenosti ali sprememb v motivaciji). Sistematični učinek lahko izkrivlja oceno ponovljivosti in povzroči, da test kljub visoki natančnosti ni stabilen v času. Tak učinek najpogosteje preverimo s parnim t-testom (v primeru dveh meritev) ali enosmerno analizo variance za ponovljene meritve (ANOVA) (v primeru treh meritev ali več). V praksi je najpogostejši razlog za tak učinek učenje in seznanjenost s testom; športnik test drugič izvede bolj samozavestno, bolje razume nalogo ali se le bolje znajde v testnem okolju. Takšna izboljšava ne odraža nujno napredka v sposobnosti, temveč le večjo izkušnost pri izvedbi. Pomembno je poudariti, da se učinek učenja ne pokaže le pri testih specifičnih veščin (npr. zahtevnejših testih ravnotežja ali koordinacijskih nalog), ampak se lahko pojavi tudi pri enostavnih testih, kot je navpični skok ali izometrično testiranje moči – še posebej, če športnik test opravlja prvič. Če

sistematični učinek odkrijemo, moramo najprej oceniti, ali gre za pričakovani in ponovljivi pojav, ki se pojavi predvsem pri prvem testiranju (npr. učni učinek, ki nato izzveni), ali za spremenljivi vpliv, kot so razlike v počitku, motivaciji ali navodilih. V prvem primeru lahko učinek omilimo z uvajalnimi ponovitvami (npr. dodatna testna seja pred začetkom spremljanja), v drugem pa je smiselno razmisliti o boljši standardizaciji postopka testiranja.

Presoja ponovljivosti in napredka na individualni ravni

Potem ko smo izbrali ustrezen test (torej takšnega, za katerega verjamemo, da je ponovljiv glede na razpoložljivo literaturo), se pojavi naslednje ključno vprašanje: ali je rezultat športnika na testu resnično boljši kot prej ali gre le za naravno variabilnost meritev? V praksi pogosto opazimo spremembo (npr. višji skok, večjo moč, hitrejši čas) in si jo pogosto takoj razlagamo kot napredek. A brez ustrezne statistične presoje ne vemo, ali ta razlika presega običajno nihanje rezultatov. Variabilnost med meritvami je povsem naravna in prisotna tudi pri zelo ponovljivih testih. Tudi pri testih, ki v literaturi izkazujejo ustrezno ponovljivost, je nujno, da spremembo rezultata posameznika ocenimo glede na stabilnost njegovih rezultatov. Za takšno presojo imamo na voljo več pristopov, ki jih lahko uporabimo pri delu s posamezniki. V nedavnem preglednem prispevku (Harry idr., 2024) se kot

uporabne omenjajo metoda statističnega modela (angl. model statistic), metoda z uporabo CV in izračun najmanjše vredne spremembe (angl. smallest worthwhile change – SWC). Vsaka od teh metod temelji na nekoliko drugačni logiki, nekatere so bolj konservativne, druge bolj občutljive. V nadaljevanju bomo predstavili prednosti in omejitve vsake metode ter prikazali, kako jih lahko uporabimo za zanesljivo presojo spremembe pri posameznem športniku, tudi če delamo z majhnimi vzorci ali posamičnimi primeri.

Za jasnejši prikaz uporabe omenjenih metod bomo razlago osvetlili s primeroma športnikov, ki sta dosegla podoben napredek v višini skoka (približno 4 cm), vendar so meritve pri prvem stabilnejše kot pri drugem. Podatki, na podlagi katerih so navedene razlage v nadaljevanju, so prikazani v tabeli 1 – izvedenih je bilo 5 testov skoka v izhodiščnem stanju in 5 testov po obdobju treninga. V praksi je smiselno, da je nabor testov še večji in da so ponovitve testa izvedene v več zaporednih dneh v obeh obdobjih.

Metoda statističnega modela

Metoda statističnega modela je pristop, ki omogoča presojo, ali je razlika med dvema povprečjema (npr. pred intervenciji in po njej) statistično značilna na ravni posameznika. Statistično je soroden t-testu za odvisne vzorce (Bates, 1996; Bates idr., 1992). Glavna ideja metode je preprosta: izračunamo razliko med povpre-

Tabela 1:

Rezultati skoka v višino pred intervenciji in po njej pri dveh športnikih z enakim napredkom, a različno ponovljivostjo

Meritev	Športnik A – pred (cm)	Športnik A – po (cm)	Športnik B – pred (cm)	Športnik B – po (cm)
1	29,8	34,2	29,0	33,5
2	30,3	33,7	33,5	32,0
3	30,1	34,1	30,5	34,5
4	29,9	34,0	27,5	35,0
5	30,0	34,4	31,0	31,0
Povprečje	30,0	34,1	30,3	34,3
SD	0,18	0,25	2,12	1,76
Razlika (cm)	+4,1		+4,0	

čjema testnih obdobj in jo primerjamo s kritično razliko – ta predstavlja najmanjšo razliko, ki je ob dani stopnji variabilnosti še lahko pripisana dejanski spremembi in ne le naključju. Če je opazovana razlika večja od kritične, sklepamo, da sprememba ni naključna. Kritična razlika se določi na podlagi povprečne standardne deviacije (SD) rezultatov obeh testnih obdobj in kritične vrednosti, ki je odvisna od števila ponovitev in izbrane stopnje tveganja (običajno $\alpha = 0,05$). Kritične vrednosti so prikazane v tabeli 2. Kritično vrednost izberemo glede na število ponovitev, ki smo jih uporabili pri posameznem testiranju, in izbrano stopnjo tveganja (najpogosteje se uporablja $\alpha = 0,05$). Če se število ponovitev med testoma razlikuje, priporočamo uporabo vrednosti, ki ustreza manjšemu številu ponovitev, saj s tem ostanemo na bolj konservativni strani presoje.

Postopek izračuna vključuje štiri korake:

1. Izračunamo absolutno razliko med povprečjema prve in druge testne seje.
2. Izračunamo povprečno standardno SD (SDPovp) obeh sej (SD_1 in SD_2) po formuli $SDPovp = \sqrt{[(SD_1^2 + SD_2^2) / 2]}$.
3. SDPovp pomnožimo s kritično vrednostjo iz ustrezne tabele (v našem primeru za 5 ponovitev in $\alpha = 0,05$ je kritična vrednost $\approx 1,366$).
4. Razliko med povprečjema primerjamo z dobljeno kritično razliko:
 - o Če je razlika večja \rightarrow sprememba statistično značilna.
 - o Če je razlika manjša ali enaka \rightarrow sprememba ni značilna.

Uporaba metode na primeru iz tabele 1

Za športnika A:

- Povprečje pred = 30,0 cm; po = 34,1 cm \rightarrow razlika = 4,1 cm.
- $SD_1 = 0,18$ cm, $SD_2 = 0,25$ cm \rightarrow SD-Povp = $\sqrt{[(0,18^2 + 0,25^2) / 2]} = 0,22$ cm.
- Kritična vrednost za 5 ponovitev pri $\alpha = 0,05 = 1,366 \rightarrow$ kritična razlika = $1,366 \times 0,22 = 0,30$ cm.
- Ker je $4,1 > 0,30 \rightarrow$ sprememba **statistično značilna**.

Za športnika B:

- Povprečje pred = 30,3 cm; po = 34,3 cm \rightarrow razlika = 4,0 cm.
- $SD_1 = 2,12$ cm, $SD_2 = 1,76$ cm \rightarrow SD-Povp = $\sqrt{[(2,12^2 + 1,76^2) / 2]} = 1,95$ cm.
- Kritična razlika = $1,366 \times 1,95 = 2,66$ cm.
- Ker je $4,0 > 2,66 \rightarrow$ sprememba **statistično značilna**.

V tem primeru bi oba športnika prestopila prag statistične značilnosti, vidimo pa, da je bil napredek pri športniku A precej večji od kritične vrednosti kot pri športniku B.

Pri uporabi metode statističnega modela moramo izbrati tudi stopnjo tveganja (α), ki predstavlja verjetnost, da zaznamo razliko, čeprav te v resnici ni (tveganje lažno pozitivnega izida). V športni praksi se najpogosteje uporablja $\alpha = 0,05$, kar pomeni, da sprejmemo 5-odstotno verjetnost, da je zaznana sprememba posledica naključja. Če želimo bolj konservativno presojo, lahko uporabimo $\alpha = 0,01$; če pa želimo biti bolj občutljivi za zaznavo manjših razlik, lahko uporabimo $\alpha = 0,10$. Višja ko je izbrana stopnja tveganja, nižja bo kritična vrednost, s tem pa bo lažje zaznati razliko, a z večjo možnostjo napačne interpretacije.

Tabela 2:

Kritične vrednosti za presojo razlike med testiranj po metodi statističnega modela

Število ponovitev	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$
3	1,3733	1,6533	2,2133
4	1,2643	1,5058	1,9867
5	1,1597	1,3662	1,7788
6	1,0629	1,2408	1,6044
7	0,9751	1,1306	1,4623
8	0,8960	1,0351	1,3473
9	0,8270	0,9536	1,2542
10	0,7673	0,8857	1,1776
11	0,7172	0,8307	1,1129
12	0,6757	0,7867	1,0581
13	0,6415	0,7516	1,0117
14	0,6132	0,7234	0,9720
15	0,5896	0,7001	0,9375
16	0,5695	0,6798	0,9070
17	0,5522	0,6618	0,8796
18	0,5371	0,6458	0,8548
19	0,5237	0,6311	0,8318
20	0,5114	0,6175	0,8102
25	0,4592	0,5572	0,7145
30	0,4194	0,5097	0,6437
35	0,3896	0,4729	0,5949
40	0,3673	0,4442	0,5626
45	0,3500	0,4207	0,5414
50	0,3352	0,4000	0,5256

Metoda koeficienta variacije

Ta pristop omogoča presojo, ali je sprememba pri posamezniku večja od variabilnosti, ki jo običajno pričakujemo pri tem testu. V nasprotju z metodo statističnega modela, ki temelji na verjetnosti, CV ne ocenjuje, ali je razlika naključna, temveč ali je dovolj velika glede na tipično nihanje rezultatov oziroma ali sprememba presega »šum« v meritvah (Turner idr., 2021). Logika metode je preprosta: če je odstotek spremembe med testnima obdobjema večji od odstotka variabilnosti v prvem testiranju, spremembo štejemo za resnično. Če pa je sprememba manjša ali primerljiva z variabilnostjo v izhodiščnem časovnem obdobju, je večja verjetnost, da gre le za običajno nihanje rezultatov. CV izračunamo kot razmerje med SD in povprečjem posameznikovega rezultata v določenem testu (najpogosteje prve testne seje), pomnoženo s 100. To nam da odstotek, ki izraža relativno velikost nihanja.

Postopek izračuna vključuje štiri korake:

1. Izračunamo povprečje in SD prve testne seje.
2. Izračunamo CV: $(SD / \text{povprečje}) \times 100$.
3. Izračunamo odstotno spremembo med obema obdobjema: $((\text{po} - \text{pred}) / \text{pred}) \times 100$.
4. Primerjamo odstotno spremembo s CV. Če je sprememba večja, gre verjetno za resnično spremembo.

Za športnika A:

- Povprečje pred: 30,0 cm; SD: 0,18 cm.
- $CV = (0,18 / 30,0) \times 100 = 0,6 \%$.
- $\text{Sprememba} = ((34,1 - 30,0) / 30,0) \times 100 = 13,7 \%$.
- Ker je $13,7 \% \gg 0,6 \%$, spremembo štejemo za resnično in pomembno.

Za športnika B:

- Povprečje pred: 30,3 cm; SD: 2,12 cm.
- $CV = (2,12 / 30,3) \times 100 = 7,0 \%$.
- $\text{Sprememba} = ((34,3 - 30,3) / 30,3) \times 100 = 13,2 \%$.
- Ker je $13,2 \% > 7,0 \%$, spremembo lahko štejemo za resnično, a z nekoliko manjšo gotovostjo.

Pri športniku A je torej sprememba jasno večja od šuma. Pri športniku B razlika prav tako presega CV, a manj izrazito. Če bi bila sprememba pri športniku B denimo le 2 cm, bi bila enaka ali celo manjša od tipične variabilnosti in sklep o napredku bi bil vprašljiv. CV je zaradi enostavne interpretacije in preračunljivosti v odstotke primeren za komunikacijo s trenerji in športniki. V kombinaciji z metodo statističnega modela omogoča celostno presojo, ali je sprememba statistično značilna in hkrati večja od šuma.

Najmanjša vredna sprememba (Smallest Worthwhile Change – SWC)

Metoda SWC poskuša odgovoriti na vprašanje, ali je sprememba ne le resnična, temveč tudi pomembna ali »vredna« v kontekstu zmogljivosti športnika. Gre za koncept, ki izhaja iz področja magnitude presoje (magnitude-based inference), pri čemer nas ne zanima le statistična značilnost, temveč tudi praktična pomembnost spremembe (Hopkins idr., 2009). Glavna ideja metode SWC je, da moramo postaviti prag, nad katerim postane sprememba funkcionalno pomembna oziroma dovolj velika, da lahko vpliva na športnikov nastop, treniranje ali selekcijo. Najpogosteje se ta prag določi kot 20 % SD rezultatov celotnega nabora podatkov (t. i. globalna SD), podobno kot po Cohenovih merilih vrednost 0,2 predstavlja standardni prag za »majhen učinek«.

Pri netreniranih posameznikih ali pri pričakovano večjih adaptacijah (na primer ob uvedbi novega trenažnega pristopa) se včasih uporablja višji prag (pogosto 60 % SD), na splošno pa je v športni diagnostiki najpogosteje uporabljen omejen prag 20 % SD.

Postopek izračuna vključuje štiri korake:

1. Združimo vse rezultate posameznika iz obeh testnih obdobji in izračunamo globalno SD (SD vseh podatkov).
2. Dobljeno SD pomnožimo z 0,2, da dobimo SWC.
3. Izračunamo razliko med povprečjema testiranj.
4. Če je razlika večja od SWC, spremembo štejemo za »vredno«; če je manjša, sprememba ni pomembna.

Uporaba metode na primeru iz tabele 1

Športnik A:

- Vseh 10 rezultatov: 29,8; 30,3; 30,1; 29,9; 30,0; 34,2; 33,7; 34,1; 34,0; 34,4.
- SD vseh rezultatov = 1,89 cm.
- $SWC = 0,2 \times 1,89 = 0,38 \text{ cm}$.
- Razlika med povprečjema = 4,1 cm → več kot desetkrat večja od SWC → sprememba je vredna.

Športnik B:

- Vseh 10 rezultatov: 29,0; 33,5; 30,5; 27,5; 31,0; 33,5; 32,0; 34,5; 35,0; 31,0.
- SD = 2,46 cm.
- $SWC = 0,2 \times 2,46 = 0,49 \text{ cm}$.
- Razlika med povprečjema = 4,0 cm → več kot osemkrat večja od SWC → sprememba je prav tako vredna.

Pri metodi SWC torej oba športnika izkazujeta pomemben napredek, vendar je treba opozoriti, da ta metoda ne upošteva variabilnosti v ponovljivosti testa (npr. individualnega šuma), temveč temelji na velikosti razlike glede na celotno variabilnost. To pomeni, da je metoda zelo občutljiva, včasih celo preveč, saj bo skoraj vedno zaznala napredek, če je le malo večji od povprečnega nihanja. Metoda SWC je uporabna kot dopolnilno orodje, zlasti kadar nas zanima praktična pomembnost sprememb. Vendar pa zaradi občutljivosti ni priporočljivo, da se uporablja samostojno. Najbolj smiselno jo je kombinirati z bolj konservativnimi merili, kot sta drugi dve opisani metodi (metoda statističnega modela in metoda CV), da ugotovimo, ali je sprememba hkrati statistično značilna in funkcionalno pomembna.

■ Zaključek

Za uspešno in zanesljivo presojo sprememb pri posamezniku je ključno, da že v izhodišču uporabljamo ustrezne teste, ki so se v literaturi izkazali kot ponovljivi (z visoko relativno in absolutno ponovljivostjo). Kazalniki, kot so ICC, SEM in CV, nam pomagajo izbrati teste, ki imajo ustrezne lastnosti na ravni skupine. Vendar pa tudi test, ki je zelo ponovljiv na ravni skupine, ne zagotavlja nujno stabilnosti rezultatov pri vsakem

posamezniku. Zato mora vsak trener poleg premišljene izbire testa na podlagi pregleda literature znati presoditi, kako ponovljiv je test v praksi za konkretnega športnika. Le z združevanjem informacij o skupinski ponovljivosti testov (kot izhodiščno vodilo) in metod za individualno presojo sprememb lahko pridemo do celovite in natančne interpretacije rezultatov. Vsaka izmed predstavljenih metod za presojo spremembe na ravni posameznika ponuja nekoliko drugačen pogled na vprašanje, ali je zaznana sprememba resnična in pomembna. Metoda statističnega modela temelji na verjetnosti in z njim presojamo, ali je razlika med testoma tako velika, da je zelo verjetno ne moremo pripisati naključju. CV ocenjuje, ali sprememba presega individualno variabilnost športnika, torej ali gre za rezultat, ki odstopa od običajnega »šuma«. SWC pa odgovarja na vprašanje, ali je sprememba dovolj velika, da bi jo lahko šteli za praktično pomembno, čeprav ne upošteva ponovljivosti meritev.

■ Literatura

- Banyard, H. G., Nosaka, K., Vernon, A. D., in Haff, G. G. (2018). The Reliability of Individualized Load–Velocity Profiles. *International Journal of Sports Physiology and Performance*, 13(6), 763–769. <https://doi.org/10.1123/ijsp.2017-0610>
- Barraza, F., Arancibia, M., Madrid, E., in Papuzinski, C. (2019). General concepts in biostatistics and clinical epidemiology: Random error and systematic error. *Medwave*, 19(07), e7687–e7687. <https://doi.org/10.5867/medwave.2019.07.7687>
- Bates, B. (1996). Single-subject methodology: an alternative approach. *Medicine & Science in Sports & Exercise*, 28(5), 631–638. <https://doi.org/10.1097/00005768-199605000-00016>
- Bates, B., Dufek, J. S., in Davis, H. P. (1992). The effect of trial size on statistical power. *Medicine and Science in Sports and Exercise*, 24(9), 1059–1065.
- Bunting, K. V., Steeds, R. P., Slater, K., Rogers, J. K., Gkoutos, G. V., in Kotecha, D. (2019). A Practical Guide to Assess the Reproducibility of Echocardiographic Measurements. *Journal of the American Society of Echocardiography*, 32(12), 1505–1515. <https://doi.org/10.1016/j.echo.2019.08.015>
- Harry, J. R., Hurwitz, J., Agnew, C., in Bishop, C. (2024). Statistical Tests for Sports Science Practitioners: Identifying Performance Gains in Individual Athletes. *Journal of Strength & Conditioning Research*, 38(5), e264–e272. <https://doi.org/10.1519/JSC.0000000000004727>
- Hopkins, W. G. (2000). Measures of Reliability in Sports Medicine and Science. *Sports Medicine*, 30(5), 375–381. <https://doi.org/10.2165/00007256-200030050-00006>
- Hopkins, W. G., Marshall, S. W., Batterham, A., in Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine and Science in Sports and Exercise*, 41(1), 3–12.
- Jiménez-Reyes, P., Samozino, P., Brughelli, M., in Morin, J. B. (2017). Effectiveness of an individualized training based on force-velocity profiling during jumping. *Frontiers in Physiology*, 7(JAN). <https://doi.org/10.3389/fphys.2016.00677>
- King, T. S., Chinchilli, V. M., in Carrasco, J. L. (2007). A repeated measures concordance correlation coefficient. *Statistics in Medicine*, 26(16), 3095–3113. <https://doi.org/10.1002/sim.2778>
- Koo, T. K., in Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–163.
- Turner, A. N., Parmar, N., Jovanovski, A., in Herne, G. (2021). Assessing Group-Based Changes in High-Performance Sport. Part 2: Effect Sizes and Embracing Uncertainty Through Confidence Intervals. *Strength & Conditioning Journal*, 43(4), 68–77. <https://doi.org/10.1519/SSC.0000000000000613>

Doc. dr. Žiga Kozinc
Fakulteta za vede o zdravju,
Univerza na Primorskem, Izola, Slovenija
ziga.kozinc@fvz.upr.si