



Perbandingan Tes Laboratorium dan Tes Manual VO₂max pada Atlet Sepak Bola Mahasiswa IKOR

Yuda El Miftah^{1*}, Agus Rusdiana¹, Tian Kurniawan¹

¹Universitas Pendidikan Indonesia, Indonesia

*Correspondence author, email: elmiftahyuda09@upi.edu

ABSTRAK

Sepak bola menuntut kapasitas daya tahan aerobik tinggi, yang diukur melalui VO₂max menggunakan tes laboratorium (CPET) sebagai baku emas maupun tes lapangan seperti Beep Test, Balke Test, dan Cooper Test. Kesesuaian ketiga tes lapangan dengan tes laboratorium pada atlet sepak bola mahasiswa Indonesia belum banyak diteliti secara bersamaan. Penelitian ini bertujuan membandingkan nilai VO₂max dari tes laboratorium dengan tiga tes lapangan pada 12 atlet sepak bola mahasiswa Program Studi Ilmu Keolahragaan (IKOR), menggunakan desain deskriptif komparatif dengan pendekatan within-subject (repeated measures). Data dianalisis dengan Repeated Measures ANOVA dan uji post hoc Bonferroni. Rata-rata VO₂max terendah diperoleh dari tes laboratorium (30,23 mL/kg/menit), sedangkan estimasi tertinggi diperoleh dari Beep Test (39,69), Balke Test (36,50), dan Cooper Test (34,59 mL/kg/menit). Repeated Measures ANOVA menunjukkan perbedaan sangat signifikan antar keempat metode ($F(1,57; 17,23) = 86,70; p < 0,001$), dan uji post hoc menunjukkan seluruh tes lapangan berbeda signifikan dengan tes laboratorium ($p < 0,05$), dengan kecenderungan mengestimasi VO₂max lebih tinggi; overestimasi terbesar pada Beep Test. Cooper Test dan Balke Test tidak berbeda signifikan. Temuan ini menegaskan ketiga tes lapangan tidak dapat menggantikan tes laboratorium tanpa kalibrasi, sehingga lebih tepat digunakan untuk pemantauan longitudinal dan interpretasi relatif, bukan nilai absolut.

Info Artikel

Diajukan: 2026-09-14

Diterima: 2026-10-13

Diterbitkan: 2026-10-14

Kata Kunci

VO₂max; tes laboratorium; CPET; beep test; balke test; cooper test; atlet sepak bola

Copyright (c) 2026 Yuda El Miftah, Agus Rusdiana, Tian Kurniawan

Jurnal KEJAORA is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License



PENDAHULUAN

Sepak bola menuntut daya tahan aerobik yang tinggi untuk mendukung aktivitas berintensitas tinggi berulang selama sekitar 90 menit pertandingan. Pemain elite menempuh sekitar 10,4 km per pertandingan pada intensitas setara 70–75% VO₂max, dan estimasi gabungan VO₂max pemain elite laki-laki sekitar 59,4 mL/kg/menit (Shalfawi & Tjelta, 2016). Karena itu, VO₂max menjadi indikator utama untuk menyaring, memantau, dan merancang beban latihan atlet. Standar acuannya adalah pengukuran langsung dengan analisis gas melalui Cardiopulmonary Exercise Testing (CPET) di treadmill (Poole & Jones, 2017), tetapi metode ini memerlukan fasilitas khusus, biaya besar, dan tenaga ahli.

Akibatnya, pelatih dan pembina olahraga di Indonesia lebih sering memakai tes lapangan, yaitu Beep Test, Balke Test, dan Cooper Test, yang mengestimasi VO₂max dari level, waktu, atau jarak tempuh melalui persamaan prediksi. Masalah praktisnya, angka hasil tes lapangan sering dibaca sebagai nilai VO₂max sebenarnya untuk menentukan kategori kebugaran dan intensitas latihan, padahal besar penyimpangannya terhadap CPET pada atlet sepak bola belum diketahui.



Bukti dari populasi lain menunjukkan bahwa penyimpangan itu nyata dan tidak seragam. Nilai prediksi Cooper Test berbeda signifikan dari pengukuran langsung pada pemain sepak bola dan hoki di India (Bandyopadhyay, 2021), tes lari bolak-balik 20 m melebihi $VO_2\text{max}$ remaja Korea sekitar 3 mL/kg/menit (Lee et al., 2019), dan 30-15 Intermittent Fitness Test melebihi $VO_2\text{max}$ pemain sepak bola mahasiswa sekitar 4,08 mL/kg/menit dengan limits of agreement yang lebar (Cheng et al., 2025). Sebaliknya, rumus Yo-Yo Endurance Test cenderung mengestimasi lebih rendah pada pemain profesional (Buglione et al., 2025). Korelasi yang tinggi, seperti pada validasi tes lapangan terhadap Astrand Treadmill Test pada mahasiswa pendidikan jasmani ($r = 0,88-0,91$; Fitriyadi, 2018), juga tidak menjamin kesesuaian nilai absolut.

Gap penelitian terlihat pada dua hal. Pertama, studi di Indonesia umumnya hanya membandingkan antar-tes lapangan tanpa CPET sebagai pembanding. Huda et al. (2026), misalnya, menemukan estimasi 26,13–34,29 mL/kg/menit pada atlet yang sama tergantung tes yang dipakai ($p = 0,001$), tetapi arah dan besar bias terhadap baku emas tidak dapat ditentukan. Kedua, studi pada pemain sepak bola yang memakai pembanding laboratorium menguji tes lain (Université de Montréal Track Test, lari 1500 m, dan Yo-Yo IR2) dan dilakukan di luar Indonesia (Düking et al., 2024; Thron et al., 2022). Sepengetahuan penulis, belum ada penelitian yang membandingkan Beep Test, Balke Test, dan Cooper Test secara bersamaan terhadap CPET pada kelompok subjek yang sama, khususnya atlet sepak bola mahasiswa.

Kebaruan penelitian ini terletak pada desain within-subject yang menempatkan ketiga tes lapangan yang paling lazim di Indonesia berhadapan langsung dengan CPET, sehingga besar dan arah bias masing-masing tes dapat dikuantifikasi dan dibandingkan. Hasilnya memberi dasar empiris bagi pelatih untuk menentukan apakah tes lapangan dapat dipakai sebagai pengganti CPET atau hanya untuk pemantauan relatif. Tujuan penelitian ini adalah membandingkan nilai $VO_2\text{max}$ dari Beep Test, Balke Test, dan Cooper Test dengan CPET pada 12 atlet sepak bola mahasiswa Program Studi Ilmu Keolahragaan (IKOR).

Berdasarkan bukti bahwa tes lapangan umumnya menyimpang dari pengukuran langsung dengan besar yang bergantung pada jenis tes dan persamaannya, hipotesis penelitian ini adalah sebagai berikut. H_0 : tidak terdapat perbedaan yang signifikan antara nilai $VO_2\text{max}$ hasil CPET dan hasil tes lapangan (Beep Test, Balke Test, dan Cooper Test) pada atlet sepak bola mahasiswa IKOR. H_1 : terdapat perbedaan yang signifikan antara nilai $VO_2\text{max}$ hasil CPET dan hasil tes lapangan pada atlet sepak bola mahasiswa IKOR, dengan besar perbedaan yang bervariasi antar tes.

METODE

Metode penelitian ini menggunakan pendekatan kuantitatif dengan desain deskriptif komparatif berpola within-subject (repeated measures), di mana peneliti tidak memberikan perlakuan apa pun, melainkan hanya membandingkan hasil pengukuran $VO_2\text{max}$ dari empat instrumen berbeda, yaitu Beep Test (Magee et al., 2021; Zaum et al., 2026), Balke Test (Fitriyadi, 2018; Huda et al., 2026), Cooper Test (Bandyopadhyay, 2021; Zaum et al., 2026), dan CPET/treadmill sebagai baku emas (Poole & Jones, 2017), pada subjek yang sama. Sampel diambil dengan teknik total sampling terhadap seluruh atlet aktif sepak bola mahasiswa Program Studi IKOR UPI yang memenuhi kriteria, sehingga terkumpul 12 orang mahasiswa laki-laki berusia 18–25 tahun yang sehat, aktif berlatih minimal tiga kali seminggu, dan telah menandatangani informed consent sesuai prinsip Deklarasi Helsinki.

Kriteria inklusi ditetapkan secara operasional sebagai berikut: (1) tercatat sebagai mahasiswa aktif Program Studi IKOR UPI yang menjadi anggota tim sepak bola program studi dan mengikuti latihan tim secara rutin, diverifikasi melalui data kemahasiswaan dan daftar hadir latihan; (2) berjenis kelamin laki-laki dan berusia 18–25 tahun, diverifikasi melalui kartu identitas; (3) berstatus sehat, yaitu tidak memiliki kontraindikasi untuk tes latihan maksimal berdasarkan pemeriksaan kesehatan awal menggunakan kuesioner PAR-Q (seluruh jawaban “tidak”), dengan tekanan darah istirahat $< 140/90$ mmHg dan denyut nadi istirahat < 100 denyut/menit; (4) aktif berlatih sepak bola minimal tiga kali per minggu dengan durasi sekitar 90 menit per sesi selama minimal enam bulan terakhir; dan (5) bersedia menjadi subjek serta telah menandatangani informed consent sesuai prinsip Deklarasi Helsinki. Kriteria eksklusi meliputi: (1) cedera muskuloskeletal atau keluhan kardiorespirasi pada saat pengambilan data; (2) mengonsumsi obat atau



suplemen yang dapat memengaruhi respons kardiorespirasi dan performa; dan (3) melakukan latihan berat dalam 24 jam sebelum tes. Subjek dikeluarkan dari analisis (drop-out) apabila tidak menyelesaikan salah satu dari keempat tes atau mengundurkan diri.

Pengukuran dilakukan selalu pada pagi hari (07.00–09.00 WIB) untuk mengontrol pengaruh variasi diurnal terhadap VO_2max , dimulai dengan pemeriksaan kesehatan awal dan pemanasan 10–15 menit guna mengurangi risiko cedera dan meningkatkan performa selama tes, lalu tes laboratorium (CPET) didahulukan sebagai baku emas, diikuti ketiga tes lapangan dengan jeda 24–48 jam antar tes (dan minimal 48 jam dari tes laboratorium) guna menghindari efek kelelahan.

Karena CPET dilaksanakan lebih dahulu sebagai baku emas, urutan ketiga tes lapangan diacak untuk setiap subjek guna mengendalikan efek urutan, yaitu kelelahan residual, familiarisasi, dan motivasi. Dengan tiga tes, tersedia enam kemungkinan urutan yang masing-masing dijalani oleh dua dari 12 subjek (counterbalanced). Pengacakan dilakukan dengan undian oleh peneliti utama sebelum pengambilan data dimulai, dan urutan setiap subjek ditetapkan sebelum subjek menjalani tes lapangan pertamanya. Hasil tes dicatat oleh penguji tanpa mengetahui urutan tes yang dijalani subjek lain.

Tes laboratorium (CPET) dilakukan di Laboratorium Sport Science Universitas Pendidikan Indonesia (UPI) pada ruangan dengan suhu $22\text{ }^{\circ}\text{C}$ dan kelembapan relatif 60%. Peralatan yang digunakan meliputi treadmill motorik dengan rentang kecepatan 0–20 km/jam dan inklinasi 0–25%, sistem analisis gas mode breath-by-breath dengan masker oronasal berukuran sesuai wajah subjek dan sensor aliran tipe turbin, serta pemantau denyut jantung berbasis sabuk dada. Sebelum setiap sesi, analisis gas dikalibrasi menggunakan gas referensi ($16,00\%$ O_2 dan $5,00\%$ CO_2) dan sensor aliran dikalibrasi dengan syringe 3 L. Protokol yang digunakan adalah protokol Bruce, diawali istirahat 3 menit dan pemanasan 5 menit, dengan peningkatan beban setiap 3 menit pada tiap tahap (kecepatan dan inklinasi dinaikkan sesuai protokol baku) hingga subjek mencapai kelelahan volunter. Nilai VO_2max ditetapkan dari rerata 30 detik konsumsi oksigen tertinggi dan dianggap valid apabila memenuhi minimal dua dari empat kriteria berikut: plateau VO_2 , rasio pertukaran respirasi $\geq 1,10$, denyut jantung $\geq 90\%$ prediksi maksimal, dan skala RPE ≥ 18 (Poole & Jones, 2017). Dorongan verbal diberikan secara standar kepada seluruh subjek.

Beep Test (multistage fitness test) dilaksanakan pada lintasan datar sepanjang 20 m yang ditandai kerucut dengan rekaman audio standar. Subjek berlari bolak-balik mengikuti bunyi "beep", mulai dari kecepatan 8,5 km/jam dengan kenaikan 0,5 km/jam setiap level (sekitar satu menit per level). Tes dihentikan ketika subjek dua kali berturut-turut gagal mencapai garis sebelum bunyi berikutnya atau berhenti karena kelelahan. Level dan shuttle terakhir dicatat, lalu dikonversi menjadi VO_2max menggunakan tabel konversi Ramsbottom et al. (1988).

Pada Cooper Test, subjek berlari atau berjalan sejauh mungkin selama 12 menit di lintasan atletik 400 m yang datar. Jarak tempuh dicatat dengan ketelitian 10 m dan VO_2max dihitung dengan persamaan $\text{VO}_2\text{max} = (\text{jarak dalam meter} - 504,9) / 44,73$ (Cooper, 1968).

Pada Balke Test, subjek berlari terus-menerus selama 15 menit di lintasan datar dengan kecepatan yang diatur sendiri. Jarak tempuh dicatat dan dikonversi menjadi VO_2max dengan persamaan $\text{VO}_2\text{max} = 33,3 + 0,178 \times (\text{jarak dalam meter} / 15 - 150)$ (Balke, 1963).

Ketiga tes lapangan dipandu oleh penguji yang sama menggunakan instruksi dan dorongan verbal yang standar, didahului pemanasan yang sama, dan dilaksanakan di Stadion UPI pada lintasan atletik berpermukaan sintetis (tartan) yang datar, dengan suhu sekitar $28\text{--}30\text{ }^{\circ}\text{C}$ dan kelembapan relatif sekitar 60–70%.

Data yang terkumpul kemudian diuji normalitasnya dengan Shapiro-Wilk, dianalisis perbedaannya antar keempat metode menggunakan Repeated Measures ANOVA (dengan koreksi Greenhouse-Geisser bila sphericity tidak terpenuhi) atau uji Friedman sebagai alternatif non-parametrik, dilanjutkan uji post hoc dengan koreksi Bonferroni untuk melihat pasangan metode yang berbeda signifikan, serta dilengkapi analisis korelasi Pearson/Spearman dan uji kesesuaian Bland-Altman untuk menilai keeratan hubungan antara tiap tes lapangan dengan tes laboratorium, dengan taraf signifikansi ditetapkan pada $\alpha = 0,05$.



HASIL DAN PEMBAHASAN

Penelitian ini melibatkan 12 atlet sepak bola mahasiswa IKOR yang menjalani empat jenis pengukuran $VO_2\text{max}$, yaitu tes laboratorium (CPET) sebagai baku emas (Poole & Jones, 2017) serta tiga tes lapangan (Beep Test, Cooper Test, dan Balke Test). Data hasil pengukuran disajikan dalam bentuk statistik deskriptif untuk memberikan gambaran umum mengenai nilai $VO_2\text{max}$ dari masing-masing metode.

1. Statistik Deskriptif $VO_2\text{max}$

Hasil perhitungan rata-rata dan simpangan baku nilai $VO_2\text{max}$ dari keempat metode pengukuran disajikan pada Tabel 1.

Tabel 1. Rata-rata dan Simpangan Baku Nilai $VO_2\text{max}$ dari Berbagai Metode Pengukuran

Metode Pengukuran	Rata-rata $VO_2\text{max}$ (mL/kg/menit)	Simpangan Baku (SD)
Tes Laboratorium (CPET)	30,23	2,04
Tes Lapangan – Beep Test	39,69	1,61
Tes Lapangan – Cooper Test	34,59	2,49
Tes Lapangan – Balke Test	36,50	2,45

Berdasarkan Tabel 1, terlihat bahwa rata-rata nilai $VO_2\text{max}$ yang diukur dengan tes laboratorium (CPET) merupakan nilai terendah di antara keempat metode, yaitu sebesar 30,23 mL/kg/menit. Sebaliknya, nilai tertinggi diperoleh dari Beep Test (39,69 mL/kg/menit), diikuti oleh Balke Test (36,50 mL/kg/menit), dan Cooper Test (34,59 mL/kg/menit). Temuan ini mengindikasikan bahwa ketiga tes lapangan cenderung memberikan estimasi $VO_2\text{max}$ yang lebih tinggi dibandingkan pengukuran langsung di laboratorium.

2. Uji Normalitas

Sebelum melakukan uji statistik parametrik, dilakukan uji normalitas menggunakan Shapiro-Wilk untuk memastikan bahwa data berdistribusi normal. Hasil uji normalitas disajikan pada Tabel 2.

Tabel 2. Hasil Uji Normalitas Shapiro-Wilk

Metode Pengukuran	Statistic	df	Sig.
CPET	0,878	12	0,083
Beep Test	0,869	12	0,064
Cooper Test	0,920	12	0,285
Balke Test	0,920	12	0,282

Hasil uji normalitas menunjukkan bahwa seluruh data dari keempat metode pengukuran memiliki nilai signifikansi (Sig.) $> 0,05$, yang berarti data berdistribusi normal. Dengan demikian, asumsi normalitas terpenuhi dan analisis statistik parametrik dapat dilanjutkan.

3. Perbandingan Antar Metode (Repeated Measures ANOVA)

Untuk menguji apakah terdapat perbedaan yang signifikan secara statistik antara keempat metode pengukuran, dilakukan uji Repeated Measures ANOVA. Uji Mauchly's Test of Sphericity menunjukkan bahwa asumsi sphericity tidak terpenuhi ($p = 0,001$), sehingga digunakan koreksi Greenhouse-Geisser. Hasil analisis menunjukkan bahwa terdapat perbedaan yang sangat signifikan antara keempat metode pengukuran $VO_2\text{max}$ ($F(1,57; 17,23) = 86,70$, $p < 0,001$). Hal ini berarti terdapat perbedaan yang bermakna pada nilai $VO_2\text{max}$ yang dihasilkan oleh tes laboratorium dengan ketiga tes lapangan.

4. Uji Post Hoc (Perbandingan Berpasangan)

Untuk mengetahui secara spesifik metode mana saja yang berbeda secara signifikan, dilakukan uji lanjutan (post hoc) dengan koreksi Bonferroni. Hasil perbandingan berpasangan disajikan pada Tabel 3.



Tabel 3. Perbandingan Berpasangan Antar Metode Pengukuran (Uji Post Hoc Bonferroni)

Metode (I)	Metode (J)	Perbedaan Rata-rata (I – J)	Sig.	Keterangan
CPET	Beep Test	9,46	0,001	Signifikan
	Cooper Test	4,36	0,002	Signifikan
	Balke Test	6,27	0,001	Signifikan
Beep Test	Cooper Test	5,10	0,003	Signifikan
	Balke Test	3,19	0,011	Signifikan
Cooper Test	Balke Test	-1,91	0,362	Tidak Signifikan

Hasil uji post hoc menunjukkan bahwa semua tes lapangan (Beep, Cooper, dan Balke) memiliki perbedaan yang signifikan terhadap hasil tes laboratorium (CPET) sebagai baku emas (Poole & Jones, 2017) ($p < 0,05$). Dengan kata lain, tidak ada satu pun dari ketiga tes lapangan yang menghasilkan nilai VO_{2max} yang setara dengan pengukuran langsung di laboratorium.

Selain itu, analisis juga mengungkapkan perbedaan yang signifikan antara hasil Beep Test dengan Cooper Test ($p = 0,003$) dan antara Beep Test dengan Balke Test ($p = 0,011$). Namun, tidak terdapat perbedaan yang signifikan antara hasil Cooper Test dan Balke Test ($p = 0,362$), yang menunjukkan bahwa kedua metode ini memiliki tingkat estimasi VO_{2max} yang relatif setara.

Penelitian ini bertujuan untuk membandingkan hasil pengukuran VO_{2max} melalui tes lapangan (Beep Test, Cooper Test, dan Balke Test) dengan tes laboratorium (CPET) pada atlet sepak bola mahasiswa IKOR. Hasil penelitian menunjukkan bahwa terdapat perbedaan yang signifikan antara pengukuran VO_{2max} melalui tes lapangan dan tes laboratorium. Temuan utama dari penelitian ini adalah bahwa semua tes lapangan secara konsisten memberikan estimasi nilai VO_{2max} yang lebih tinggi dibandingkan pengukuran langsung menggunakan CPET/treadmill sebagai baku emas (Poole & Jones, 2017).

Perbedaan yang paling mencolok terlihat pada Beep Test, yang menunjukkan rata-rata overestimasi sebesar 9,46 mL/kg/menit, diikuti oleh Balke Test (+6,27 mL/kg/menit), dan Cooper Test (+4,36 mL/kg/menit). Overestimasi ini dapat dijelaskan oleh beberapa faktor. Pertama, tes lapangan pada dasarnya adalah metode estimasi tidak langsung. Nilai VO_{2max} diprediksi berdasarkan kinerja fisik (seperti jarak, waktu, atau level yang dicapai) dengan menggunakan rumus atau tabel prediksi yang dikembangkan dari populasi tertentu. Rumus-rumus ini, seperti yang digunakan dalam Cooper Test dan Balke Test, mungkin tidak sepenuhnya berlaku untuk karakteristik fisik, tingkat kebugaran, atau efisiensi gerak atlet sepak bola dalam penelitian ini (Bandyopadhyay, 2021; Huda et al., 2026).

Kedua, faktor lingkungan juga berperan. Tes lapangan umumnya dilakukan di luar ruangan dengan kondisi yang kurang terkontrol (misalnya, suhu, kelembaban, dan permukaan lapangan), yang dapat memengaruhi performa subjek. Sebaliknya, tes laboratorium dilakukan dalam lingkungan yang terkontrol, memungkinkan pengukuran yang lebih stabil dan akurat. Ketiga, aspek psikologis dan motivasi juga menjadi faktor. Dalam tes lapangan seperti Cooper Test dan Beep Test, subjek harus mengatur sendiri kecepatan dan strategi larnya. Hal ini dapat menyebabkan variasi performa yang lebih besar dan tidak selalu mencerminkan upaya maksimal yang sebenarnya, sehingga dapat merendahkan atau meninggikan hasil estimasi. Tes laboratorium dengan protokol yang meningkat secara bertahap memungkinkan pengawasan langsung dan dorongan verbal untuk memastikan subjek mencapai kelelahan maksimal (Poole & Jones, 2017).

Hasil ini sejalan dengan Thron et al. (2022) yang menemukan bahwa tes lapangan pada pemain sepak bola terlatih melebihi kecepatan aerobik maksimal hasil laboratorium sebesar 1,61–1,68 km/jam, meskipun parameter dan protokol yang digunakan berbeda. Düking et al. (2024) juga melaporkan kesesuaian yang buruk antara Yo-Yo IR2 dan analisis gas langsung pada pemain muda elite ($ICC = 0,17$), meskipun besar dan arah selisihnya berbeda dari penelitian ini. Di Indonesia, Huda et al. (2026) mendapati bahwa estimasi VO_{2max} dari Multistage Fitness Test, Rockport, Cooper, dan Balke berbeda signifikan pada atlet Tarung Derajat ($p = 0,001$), dengan Cooper Test menghasilkan nilai terendah, berlawanan dengan penelitian ini yang Cooper Test-nya mengestimasi lebih tinggi daripada CPET, sehingga arah dan besar



bias tes lapangan tampaknya bergantung pada populasi dan protokol. Pada populasi Asia, Bandyopadhyay (2021) juga menemukan bahwa nilai prediksi Cooper Test berbeda signifikan dari pengukuran langsung pada pemain sepak bola dan hoki di Kolkata, India, sehingga ketidaksesuaian tes lapangan terhadap baku emas tidak hanya ditemukan pada penelitian ini. Temuan serupa dilaporkan pada remaja Korea, yang VO_{2max} dari tes lari bolak-balik 20 m termodifikasi-nya lebih tinggi sekitar 2,95 mL/kg/menit daripada treadmill (Lee et al., 2019). Sebaliknya, pada pemain profesional, rumus estimasi Yo-Yo Endurance Test justru cenderung mengestimasi VO_{2max} lebih rendah daripada pengukuran langsung (Buglione et al., 2025), sehingga arah bias bergantung pada jenis tes dan persamaan yang digunakan. Hasil ini juga mendukung pernyataan Poole dan Jones (2017) yang menegaskan bahwa pengukuran langsung melalui CPET merupakan metode acuan dalam menentukan VO_{2max} .

Perbandingan dengan studi lain menunjukkan bahwa besar bias bergantung pada jenis tes, persamaan prediksi, dan karakteristik subjek. Magee et al. (2021) pada 65 atlet hoki lapangan putri perguruan tinggi (VO_{2max} laboratorium $46,4 \pm 4,6$ mL/kg/menit) menemukan bahwa persamaan Ramsbottom ($46,5$ mL/kg/menit) dan Flouris ($46,3$ mL/kg/menit) menghasilkan estimasi Beep Test yang valid, sedangkan beberapa persamaan lain melebihi VO_{2max} , dan mereka menekankan bahwa usia, tingkat kebugaran, dan riwayat latihan perlu dipertimbangkan dalam memilih persamaan. Kumar (2021) juga melaporkan bahwa seluruh nilai prediksi multistage shuttle run test melebihi hasil treadmill, meskipun hanya pada empat subjek. Overestimasi Beep Test pada penelitian ini ($9,46$ mL/kg/menit) melampaui rentang kesalahan baku estimasi (SEE) sebesar $2,5-4,1$ mL/kg/menit yang dilaporkan dalam kajian Zaum et al. (2026) pada remaja, sedangkan bias Cooper Test ($4,36$ mL/kg/menit) berada dalam rentang SEE Cooper Test mereka ($3,8-5,6$ mL/kg/menit). Dengan demikian, selisih pada Beep Test di kelompok ini tidak dapat dianggap semata-mata sebagai galat acak. Di sisi lain, korelasi antara tes lapangan dan tes laboratorium dapat tetap tinggi walaupun nilai absolutnya bias, sebagaimana dilaporkan Thron et al. (2022) (r sekitar $0,65-0,87$), sehingga tes lapangan masih layak digunakan untuk memeringkat atau memantau atlet secara relatif. Pada pemain sepak bola mahasiswa laki-laki, Cheng et al. (2025) melaporkan bahwa 30-15 Intermittent Fitness Test berkorelasi besar hingga sangat besar dengan VO_{2max} laboratorium ($r = 0,62-0,77$) tetapi melebihi nilainya rata-rata sekitar $4,08$ mL/kg/menit dengan limits of agreement yang lebar, sehingga korelasi yang kuat tidak menjamin kesesuaian nilai absolut, searah dengan overestimasi pada penelitian ini. Tinjauan sistematis terhadap 248 studi tes Yo-Yo juga menyimpulkan bahwa bukti validitas terhadap VO_{2max} laboratorium bersifat campuran dan variabilitas tes-ulangnya berkisar $4,5-12,7\%$, sehingga tes lapangan perlu ditafsirkan hati-hati untuk perubahan yang kecil (Schmitz et al., 2018).

Menariknya, meskipun semua tes lapangan berbeda secara signifikan dengan CPET, Cooper Test dan Balke Test menunjukkan hasil yang tidak berbeda secara signifikan satu sama lain. Ini mengindikasikan bahwa kedua formula estimasi (berdasarkan jarak tempuh 12 menit dan waktu tempuh hingga kelelahan) mungkin memiliki tingkat akurasi yang serupa untuk populasi ini. Namun, keduanya tetap lebih rendah tingkat kesesuaiannya dibandingkan Beep Test yang memiliki overestimasi paling besar. Hal ini mungkin disebabkan oleh sifat Beep Test yang bersifat maximal intermittent, di mana subjek harus melakukan akselerasi dan deselerasi berulang kali, sebuah aktivitas yang kurang efisien secara metabolik dan fisiologis dibandingkan lari kontinu di atas treadmill, sehingga prediksi VO_{2max} -nya menjadi lebih tinggi.

Secara fisiologis, VO_{2max} ditentukan oleh rantai transpor dan pemanfaatan oksigen, yaitu curah jantung, kapasitas difusi paru, kapasitas angkut oksigen darah, distribusi aliran darah ke otot rangka, dan kapasitas mitokondria otot untuk menggunakan oksigen, sebagaimana dirumuskan dalam prinsip Fick ($VO_2 = \text{curah jantung} \times \text{perbedaan kandungan oksigen arteri-vena}$) (Poole & Jones, 2017). CPET mengukur hasil akhir seluruh rantai tersebut secara langsung dari selisih konsentrasi oksigen udara inspirasi dan ekspirasi yang dikalikan ventilasi menit. Sebaliknya, tes lapangan tidak mengukur satu pun komponen rantai tersebut, melainkan memprediksi VO_{2max} dari performa lari (jarak, waktu, atau tahap) yang juga dipengaruhi oleh ekonomi lari, kapasitas anaerob, keterampilan lari, dan strategi pengaturan kecepatan.

Pada Beep Test, subjek harus berakselerasi, melambat, dan berbalik arah setiap 20 m dengan kecepatan yang meningkat tiap tahap. Lari tak-tunak semacam ini diduga membutuhkan energi metabolik per satuan jarak yang lebih besar daripada lari berkecepatan konstan, dan pada tahap akhir tes sebagian



energinya dipasok oleh sistem anaerob. Persamaan prediksi menerjemahkan tahap terakhir yang dicapai menjadi VO_{2max} seolah-olah seluruhnya ditopang metabolisme aerobik, sehingga kontribusi anaerob dan keterampilan berbalik arah, yang pada pemain sepak bola relatif terlatih, ikut terbaca sebagai kapasitas aerobik dan menghasilkan overestimasi terbesar. Sebaliknya, CPET menggunakan lari kontinu dengan peningkatan beban bertahap sehingga VO_2 mendekati nilai maksimal terutama karena keterbatasan sistem kardiorespirasi dan otot.

Persamaan Cooper dan Balke berbentuk regresi linear yang menghubungkan jarak tempuh dengan VO_{2max} pada populasi asalnya, dengan asumsi bahwa ekonomi lari dan fraksi VO_{2max} yang dapat dipertahankan relatif seragam antar individu. Apabila subjek memiliki ekonomi lari yang lebih baik, atau menempuh jarak yang sama dengan kontribusi anaerob yang lebih besar, jarak yang sama akan diterjemahkan menjadi nilai VO_{2max} yang lebih tinggi daripada konsumsi oksigen sebenarnya. Perbedaan permukaan dan kondisi lingkungan, misalnya hambatan udara dan suhu di lapangan dibandingkan treadmill di ruangan terkontrol, juga dapat mengubah biaya energi lari. Selain itu, pada subjek dengan VO_{2max} laboratorium yang relatif rendah seperti pada penelitian ini (30,23 mL/kg/menit), variasi kecil pada keterampilan lari dan pengaturan kecepatan dapat menghasilkan selisih estimasi yang proporsional lebih besar. Penjelasan mekanisme ini bersifat inferensial karena penelitian ini tidak mengukur ekonomi lari, kontribusi anaerob, atau konsentrasi laktat, sehingga perlu diuji pada penelitian selanjutnya.

Perbedaan ini memiliki implikasi penting bagi pelatih dan staf pendukung atlet. Penggunaan tes lapangan secara tunggal tanpa kalibrasi dapat memberikan gambaran yang keliru tentang kapasitas aerobik atlet. Jika pelatih hanya mengandalkan hasil Beep Test yang tinggi, mereka mungkin meyakini bahwa atlet memiliki daya tahan aerobik yang lebih baik dari kenyataannya, yang dapat memengaruhi perencanaan dan evaluasi program latihan. Oleh karena itu, penting untuk menyadari bahwa nilai VO_{2max} dari tes lapangan adalah estimasi, bukan nilai absolut. Untuk evaluasi yang paling akurat, terutama dalam konteks ilmiah atau pembinaan atlet elite, tes laboratorium tetap menjadi pilihan terbaik (Poole & Jones, 2017). Pada pemain sepak bola profesional, CPET treadmill dengan analisis gas napas-per-napas dapat mendeteksi penurunan VO_{2max} dengan ukuran efek sedang ($d = 0,61$) setelah infeksi COVID-19 (Parpa & Michaelides, 2022), suatu kepekaan yang sulit dipenuhi oleh tes lapangan dengan bias sebesar yang ditemukan pada penelitian ini.

Meskipun demikian, tes lapangan tetap memiliki peran penting dalam praktik olahraga karena kepraktisannya. Untuk meningkatkan kegunaannya, penelitian lebih lanjut disarankan untuk mengembangkan rumus prediksi spesifik untuk atlet sepak bola Indonesia, atau menggunakan koreksi faktor terhadap hasil tes lapangan berdasarkan perbedaan yang ditemukan dalam penelitian ini. Dengan demikian, tes lapangan dapat tetap menjadi alat yang berharga untuk memantau perkembangan kebugaran atlet secara berkala dan murah, dengan tetap menyadari keterbatasannya. Perangkat wearable yang divalidasi terhadap CPET treadmill juga mulai dipertimbangkan sebagai alternatif, tetapi validitasnya bergantung pada populasi sehingga perlu diuji pada atlet sepak bola mahasiswa (Železnik Mežan, 2025).

Keterbatasan penelitian ini perlu dipertimbangkan dalam menafsirkan hasil. Pertama, jumlah sampel kecil ($n = 12$), berasal dari satu program studi, dan seluruhnya laki-laki, sehingga daya statistik dan generalisasi temuan terbatas, termasuk terhadap pemain perempuan, pemain elite, maupun posisi bermain tertentu. Nilai VO_{2max} laboratorium rerata (30,23 mL/kg/menit) tergolong relatif rendah untuk atlet sepak bola, jauh di bawah estimasi gabungan sekitar 59,4 mL/kg/menit pada pemain sepak bola elite laki-laki (Shalfawi & Tjelta, 2016) maupun nilai sekitar 62 mL/kg/menit pada pemain profesional yang diuji dengan CPET treadmill (Silva et al., 2018), sedangkan tingkat kebugaran memengaruhi akurasi persamaan prediksi tes lapangan (Magee et al., 2021); arah dan besar bias pada atlet dengan VO_{2max} lebih tinggi dapat berbeda. Kedua, urutan tes tetap (CPET lebih dahulu, kemudian tiga tes lapangan) dan tidak diacak atau diseimbangkan (counterbalanced), sehingga efek urutan, pembelajaran, dan kelelahan kumulatif tidak dapat dikesampingkan sepenuhnya walaupun jeda 24–48 jam telah diberikan. Ketiga, kondisi lingkungan pada tes lapangan tidak dapat dikendalikan seketat pada tes laboratorium, dan upaya maksimal pada tes lapangan bergantung pada motivasi serta strategi pengaturan kecepatan subjek. Keempat, estimasi VO_{2max} tes lapangan dihitung dari persamaan prediksi tertentu yang dikembangkan pada populasi lain, sedangkan



penelitian ini tidak membandingkan persamaan alternatif, padahal pemilihan persamaan dapat mengubah besar bias (Magee et al., 2021). Kelima, setiap tes dilakukan dengan desain deskriptif komparatif tanpa pengukuran ulang, sehingga reliabilitas tes-ulang masing-masing metode tidak dievaluasi, dan mekanisme fisiologis yang dibahas bersifat inferensial. Penelitian selanjutnya disarankan menggunakan sampel yang lebih besar dan multi-kampus, desain dengan urutan tes yang diseimbangkan, pengukuran ekonomi lari dan laktat, serta pengembangan persamaan koreksi atau prediksi khusus untuk atlet sepak bola Indonesia.

KESIMPULAN

Pada 12 atlet sepak bola mahasiswa laki-laki Program Studi IKOR ($VO_2\text{max}$ laboratorium rerata 30,23 mL/kg/menit) yang diukur satu kali dengan desain within-subject, nilai $VO_2\text{max}$ dari Beep Test, Balke Test, dan Cooper Test berbeda signifikan dari hasil CPET (Repeated Measures ANOVA dengan koreksi Greenhouse-Geisser, $p < 0,001$; post hoc Bonferroni, seluruh $p < 0,05$), sehingga H_0 ditolak. Rerata ketiga tes lapangan lebih tinggi daripada CPET, dengan selisih terbesar pada Beep Test (+9,46 mL/kg/menit), diikuti Balke Test (+6,27 mL/kg/menit) dan Cooper Test (+4,36 mL/kg/menit). Cooper Test dan Balke Test tidak berbeda signifikan satu sama lain ($p = 0,362$), tetapi hasil ini tidak membuktikan keduanya setara karena sampel kecil dan tidak dilakukan uji ekuivalensi. Selisih yang dilaporkan adalah selisih rerata kelompok, sehingga kesesuaian pada tingkat individu tidak disimpulkan.

Dalam batas data ini, nilai estimasi tes lapangan sebaiknya tidak dibaca sebagai nilai $VO_2\text{max}$ absolut atau dipakai menggantikan CPET tanpa kalibrasi. Temuan ini berlaku untuk kelompok, tes, dan persamaan prediksi yang digunakan, serta belum dapat digeneralisasikan ke atlet perempuan, atlet elite atau berkebugaran lebih tinggi, maupun populasi sepak bola lain, mengingat sampel kecil, berasal dari satu program studi, dan memiliki $VO_2\text{max}$ laboratorium yang relatif rendah. Penelitian ini juga tidak menguji reliabilitas tes-ulang maupun kepekaan tes terhadap perubahan, sehingga kelayakan tes lapangan untuk pemantauan longitudinal belum dapat disimpulkan dan perlu diuji pada sampel yang lebih besar dengan pengukuran berulang dan analisis kesesuaian individual.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Pendidikan Indonesia dan Program Studi Ilmu Keolahragaan, Fakultas Pendidikan Olahraga dan Kesehatan, atas dukungan fasilitas, sarana, dan izin yang diberikan selama proses penelitian ini berlangsung. Penulis juga menyampaikan terima kasih kepada Agus Rusdiana dan Tian Kurniawan selaku dosen pembimbing atas arahan, bimbingan, dan masukan yang sangat berharga sepanjang proses penelitian ini. Ucapan terima kasih juga disampaikan kepada seluruh subjek penelitian, yaitu mahasiswa atlet sepak bola Program Studi Ilmu Keolahragaan, atas kesediaan, waktu, dan partisipasi aktifnya dalam mengikuti seluruh rangkaian tes pengukuran $VO_2\text{max}$, sehingga penelitian ini dapat terlaksana dengan baik.

DAFTAR PUSTAKA

- Bandyopadhyay, A. (2021). Validity of Cooper's 12-min run test for estimation of maximum oxygen uptake in hockey and soccer players of Kolkata, India. *Journal of Human Ergology*, 50(1), 39–43. https://doi.org/10.11183/jhe.50.1_39
- Buglione, A., Pompa, D., Beato, M., Rocchi, M. B. L., Savoia, C., Bertollo, M., Curzi, D., Sisti, D., & Perroni, F. (2025). Measured vs. estimated $VO_2\text{max}$ in the Yo-Yo Endurance Test: An exploratory study in professional soccer players. *Sports*, 13(12), Article 424. <https://doi.org/10.3390/sports13120424>
- Cheng, R., Lin, W., Song, L., Pan, J., Wang, N., & Li, X. (2025). Re-examining the reliability and validity of 30-15IFT for $VO_2\text{max}$ prediction in male collegiate soccer players: A pilot study. *Frontiers in Physiology*, 16, Article 1668250. <https://doi.org/10.3389/fphys.2025.1668250>
- Düking, P., Ruf, L., Altmann, S., Thron, M., Kunz, P., & Sperlich, B. (2024). Assessment of maximum oxygen uptake in elite youth soccer players: A comparative analysis of smartwatch technology, Yoyo Intermittent Recovery Test 2, and respiratory gas analysis. *Journal of Sports Science and Medicine*, 23, 351–357. <https://doi.org/10.52082/jssm.2024.351>



- Fitriyadi, G. (2018). Perbandingan validitas tes $VO_2\text{max}$ antara metode maksimal dan sub-maksimal pada remaja. *Gelombang Pendidikan Jasmani Indonesia*, 2(2), 116–119. <https://doi.org/10.17977/um040v2i2p116-119>
- Huda, N., Hadiono, Kristi, P. D., & Kustiawan, A. A. (2026). Analisis komparatif estimasi $VO_2\text{max}$ antara tes multistage, Rockport, Cooper, dan Balke. *Indonesian Journal of Sport Science and Technology*, 5(1), 684–690. <https://doi.org/10.31316/ijst.v5i1.9826>
- Kumar, A. S. (2021). Validity of the 20 meter multi-stage shuttle run test (MSSRT) in predicting $VO_2\text{max}$. *ACPES Journal of Physical Education, Sport, and Health*, 1(2), 131–133. <https://doi.org/10.15294/ajpesh.v1i2.52587>
- Lee, S.-H., Song, J.-R., Kim, Y.-J., Kim, S.-J., Park, H., Kim, C.-S., Kwak, H.-B., Kang, J.-H., & Park, D.-H. (2019). New 20 m progressive shuttle test protocol and equation for predicting the maximal oxygen uptake of Korean adolescents aged 13–18 years. *International Journal of Environmental Research and Public Health*, 16(13), Article 2265. <https://doi.org/10.3390/ijerph16132265>
- Magee, M. K., White, J. B., Merrigan, J. J., & Jones, M. T. (2021). Does the multistage 20-m shuttle run test accurately predict $VO_2\text{max}$ in NCAA Division I women collegiate field hockey athletes? *Sports*, 9(6), Article 75. <https://doi.org/10.3390/sports9060075>
- Parpa, K., & Michaelides, M. (2022). Aerobic capacity of professional soccer players before and after COVID-19 infection. *Scientific Reports*, 12, Article 11850. <https://doi.org/10.1038/s41598-022-16031-7>
- Poole, D. C., & Jones, A. M. (2017). Measurement of the maximum oxygen uptake $VO_2\text{max}$: $VO_2\text{peak}$ is no longer acceptable. *Journal of Applied Physiology*, 122(4), 997–1002. <https://doi.org/10.1152/jappphysiol.01063.2016>
- Schmitz, B., Pfeifer, C., Kreitz, K., Borowski, M., Faldum, A., & Brand, S.-M. (2018). The Yo-Yo intermittent tests: A systematic review and structured compendium of test results. *Frontiers in Physiology*, 9, Article 870. <https://doi.org/10.3389/fphys.2018.00870>
- Shalfawi, S. A. I., & Tjelta, L. I. (2016). A critical evaluation of the aerobic capacity demands of elite male soccer players. *International Journal of Applied Sports Sciences*, 28(2), 200–212. <https://ijass.sports.re.kr/v.28/2/200/836>
- Silva, C. G. S., Castro, C. L. B., Franca, J. F., Bottino, A., Myers, J., & Araújo, C. G. S. (2018). Cardiorespiratory optimal point in professional soccer players: A novel submaximal variable during exercise. *International Journal of Cardiovascular Sciences*, 31(4), 323–332. <https://doi.org/10.5935/2359-4802.20180030>
- Thron, M., Woll, A., Klos, L., Härtel, S., Ruf, L., Kloss, C., & Altmann, S. (2022). Overestimation of maximal aerobic speed by the Université de Montréal track test and a 1500-m-time trial in soccer. *Frontiers in Physiology*, 13, Article 1023257. <https://doi.org/10.3389/fphys.2022.1023257>
- Zaum, S., Ridjal, A. T. M., & Hasanah, U. (2026). Comparison of the accuracy of the bleep test, Yo-Yo IR, and Cooper test for estimating $VO_2\text{max}$ in adolescents: A literature review. *COMPETITOR: Jurnal Pendidikan Kepeleatihan Olahraga*, 18(2), 4439–4450. <https://doi.org/10.26858/cpjok.v18i2.981>
- Železnik Mežan, L. (2025). Accuracy of wearables for determining the maximal oxygen uptake and lactate threshold: A qualitative systematic review. *Frontiers in Sports and Active Living*, 7, Article 1707991. <https://doi.org/10.3389/fspor.2025.1707991>