KOLONI: Jurnal Multidisiplin Ilmu, 5 (4), Tahun 2026
e-ISSN: 2828-6863

Validitas Dan Reliabilitas *Logbook* Digital Perilaku Sedentari, Neat, Dan Eat Mahasiswa Kedokteran

Aisyah Dita Shafira[1]*, Esa Indah Ayudia Tan[2], Patrick William Gading[3], Huntari Harahap[4]
[1,2,3,4]Program Studi Kedokteran, Fakultas Kedokteran dan Ilmu Kesehatan, Universitas Jambi, Jambi, Indonesia
Corresponding author: 05ditaaisyah@gmail.com

Abstrak: Penelitian ini menilai validitas isi, validitas muka, dan reliabilitas *test-retest logbook* digital berbasis *ecological momentary assessment* yang mengukur *sedentary behavior* (SB), *non-exercise activity thermogenesis* (NEAT), dan *exercise activity thermogenesis* (EAT) secara terpisah. *Logbook* dengan nilai MET dari *2024 Adult Compendium of Physical Activities* dinilai oleh dua pakar *(content validity index*, CVI) dan diuji pada 20 mahasiswa kedokteran selama dua minggu berturut-turut. Reliabilitas dianalisis dengan *intraclass correlation coefficient* (ICC) *two-way mixed, absolute agreement* (kriteria ≥0,70). Seluruh 52 item yang dinilai pakar memperoleh I-CVI 1,00, dan versi akhir berisi 53 item setelah satu item dipecah. S-FVI/Ave sebesar 0,97. Nilai ICC SB 0,491 (IK95% 0,065-0,764), NEAT 0,720 (0,363-0,884), dan EAT 0,898 (0,763-0,958). Skor NEAT pada minggu kedua lebih rendah secara bermakna (p=0,019), dan ICC EAT turun menjadi 0,346 setelah satu responden ekstrem dikeluarkan. Validitas isi dan muka sangat baik, tetapi reliabilitas bervariasi. Instrumen lebih tepat digunakan untuk analisis kelompok dan perlu penyempurnaan.

Kata kunci: aktivitas fisik, *logbook*, perilaku sedentari, reliabilitas, validitas

Abstract: *This study assessed content validity, face validity, and test-retest reliability of an ecological momentary assessment-based digital logbook that measures sedentary behavior (SB), non-exercise activity thermogenesis (NEAT), and exercise activity thermogenesis (EAT) separately. A logbook with MET values from the 2024 Adult Compendium of Physical Activities was rated by two experts (content validity index, CVI) and tested in 20 medical students over two consecutive weeks. Reliability was analyzed using intraclass correlation coefficient (ICC), two-way mixed, absolute agreement (criterion ≥0.70). All 52 rated items achieved I-CVI 1.00, and one item was then split, giving 53 final items. S-FVI/Ave was 0.97. ICC was 0.491 (95% CI 0.065-0.764) for SB, 0.720 (0.363-0.884) for NEAT, and 0.898 (0.763-0.958) for EAT. Week-2 NEAT scores were significantly lower (p=0.019), and EAT ICC fell to 0.346 after excluding one extreme respondent. Content and face validity were excellent, but reliability varied. The instrument suits group-level analysis but needs refinement.*

Keywords: *physical activity, logbook, sedentary behavior, reliability, validity*

Pendahuluan

Prevalensi *overweight* dan obesitas yang terus meningkat menjadi masalah kesehatan global. Pada tahun 2024, sebanyak 2,6 miliar orang dewasa mengalami *overweight*, dan 940 juta di antaranya mengalami obesitas (World Health Organization, 2026). Bahkan, pada mahasiswa kedokteran Universitas Jambi angkatan 2022-2024, prevalensi *overweight* dan obesitas pada 200 responden mencapai 49% (Yulinda et al., 2026). Hal ini mengkhawatirkan karena masa kuliah merupakan periode yang penting, ketika kebiasaan gaya hidup yang terbentuk cenderung terbawa hingga dewasa (Telleria-Aramburu & Arroyo-Izaga, 2022). Salah satu gaya hidup yang dikhawatirkan terbentuk pada periode ini adalah perilaku sedentari, karena perilaku sedentari yang tinggi dilaporkan sebagai salah satu faktor risiko gizi lebih pada dewasa muda (Setyaningsih & Fatqur, 2025). Namun, hubungan perilaku sedentari dengan indeks massa tubuh (IMT) pada mahasiswa belum konsisten. Satu penelitian menemukan adanya hubungan (Yuniarti et al., 2024), tetapi penelitian lain pada mahasiswa tingkat akhir tidak menemukan hubungan tersebut (Putri et al., 2026). Perilaku sedentari juga diduga berkontribusi terhadap rendahnya *physical*

activity energy expenditure (PAEE), yaitu komponen *total energy expenditure* (TEE) yang paling bervariasi antarindividu (Bonilla et al., 2023; Chung et al., 2018; von Loeffelholz & Birkenfeld, 2022).

PAEE terdiri atas dua komponen, yaitu *exercise activity thermogenesis* (EAT) dan *non-exercise activity thermogenesis* (NEAT). EAT adalah pengeluaran energi dari olahraga (Bonilla et al., 2023; Rizzato et al., 2022), yaitu aktivitas fisik yang terencana, terstruktur, dan berulang untuk menjaga atau meningkatkan kebugaran (Caspersen et al., 1985). Sementara itu, NEAT adalah pengeluaran energi dari aktivitas spontan di luar olahraga, seperti berdiri, berjalan kaki, dan mempertahankan postur (Rizzato et al., 2022; von Loeffelholz & Birkenfeld, 2022). NEAT dapat berbeda hingga 2.000 kkal per hari antarindividu dan cenderung rendah pada individu dengan obesitas (Bonilla et al., 2023). Kedua komponen ini berbeda dari *sedentary behavior* (SB), yaitu perilaku saat terjaga dengan pengeluaran energi ≤1,5 *metabolic equivalent of task* (MET) dalam posisi duduk, bersandar, atau berbaring (Tremblay et al., 2017). Pada masa pandemi COVID-19, proporsi mahasiswa Indonesia yang memenuhi rekomendasi aktivitas fisik dilaporkan paling rendah (45%) di antara tujuh negara ASEAN yang diteliti (Amornsriwatanakul et al., 2022). Untuk mengetahui bagian aktivitas mana yang kurang, SB, NEAT, dan EAT perlu diukur secara terpisah.

Namun, instrumen yang biasa dipakai pada mahasiswa, seperti bagian *sitting time* dari *International Physical Activity Questionnaire* (IPAQ), umumnya hanya mengukur SB sebagai total waktu duduk dengan validitas kriteria yang rendah (Bakker et al., 2020; Meh et al., 2021). Instrumen tersebut juga tidak dirancang untuk memisahkan NEAT dan EAT. Bahkan, pada mahasiswa ilmu keolahragaan di Indonesia, IPAQ versi pendek tidak berhubungan bermakna dengan akselerometer pada berbagai intensitas aktivitas fisik (Akbar et al., 2025). Akselerometer memang lebih objektif, tetapi tidak dapat menangkap jenis dan konteks SB seperti metode swalapor (Bakker et al., 2020). Sebaliknya, metode log atau *diary* memiliki validitas kriteria yang lebih tinggi dibandingkan kuesioner dalam mengukur SB, yaitu r=0,63 dibandingkan dengan 0,35 (Bakker et al., 2020). Jika pencatatan dilakukan beberapa kali sehari dengan pendekatan *ecological momentary assessment* (EMA), bias mengingat juga dapat dikurangi karena perilaku dicatat tidak lama setelah dilakukan (Hartson et al., 2023; S. Wang et al., 2025).

Peneliti mengembangkan *logbook* aktivitas fisik digital berbasis EMA yang berisi daftar aktivitas SB, NEAT, dan EAT yang biasa dilakukan mahasiswa, dengan nilai MET dari *2024 Adult Compendium of Physical Activities* (Herrmann et al., 2024). *Logbook* ini akan digunakan pada penelitian hubungan SB, NEAT, dan EAT dengan status *overweight*/obesitas pada mahasiswa, sehingga perlu mencatat jenis aktivitas, tidak hanya total durasinya. Kuesioner NEAT berbasis *Compendium* memang pernah dikembangkan dan divalidasi pada pasien diabetes melitus tipe 2 di Jepang, tetapi hanya mengukur NEAT dengan skor ordinal (Hamasaki et al., 2014). Oleh

karena itu, instrumen yang umum dipakai belum memisahkan NEAT dan EAT, dan kuesioner NEAT yang tersedia belum ditujukan untuk mahasiswa. Sejauh penelusuran peneliti di Garuda, Google Scholar, dan PubMed hingga Oktober 2026, juga belum ada instrumen berbahasa Indonesia yang sudah diuji validitas dan reliabilitasnya untuk mengukur ketiga komponen tersebut secara terpisah pada mahasiswa. Penelitian ini bertujuan menilai validitas isi, validitas muka, dan reliabilitas *test-retest logbook* tersebut pada mahasiswa Program Studi Kedokteran, Fakultas Kedokteran dan Ilmu Kesehatan (FKIK) Universitas Jambi.

Metode

Penelitian kuantitatif ini merupakan studi validasi instrumen untuk penelitian potong lintang mengenai hubungan SB, NEAT, dan EAT dengan status *overweight*/obesitas pada mahasiswa FKIK Universitas Jambi. Penelitian dilaksanakan di FKIK Universitas Jambi pada Agustus-September 2026 dan telah disetujui oleh Komite Etik Penelitian Kesehatan FKIK Universitas Jambi. Seluruh responden telah memberikan persetujuan setelah penjelasan (*informed consent*). *Logbook* disusun pada platform KoBoToolbox. Versi yang dinilai pakar berisi 52 item. Atas masukan salah satu pakar, item *high-intensity interval training* (HIIT)/*calisthenics* keras dipecah menjadi dua, dengan kode MET yang diusulkan peneliti dan disetujui oleh pakar tersebut. Versi akhir berisi 53 item (11 item SB, 16 item NEAT, dan 26 item EAT) dan sudah digunakan sejak uji coba minggu pertama. Setiap item dilengkapi kode aktivitas dan nilai MET dari *Compendium* (Herrmann et al., 2024), dan setiap kategori menyediakan kolom isian untuk aktivitas lain. Responden mencentang aktivitas dan mengisi durasinya dalam menit pada tiga sesi per hari, yaitu bangun tidur-13.00, 13.00-17.00, dan 17.00-21.00, dengan pengingat melalui WhatsApp. Aktivitas pukul 21.00 hingga tidur ditanyakan pada sesi 1 keesokan harinya sebagai satu aktivitas dominan beserta durasinya. Jam tidur, jam bangun, dan tidur sebentar (*nap*) juga ditanyakan untuk memeriksa apakah total durasi masih berada dalam waktu terjaga. *Logbook* diisi selama tiga hari, yaitu satu hari libur (*weekend*) dan dua hari perkuliahan (*weekday*).

Validitas isi dinilai oleh dua pakar, yaitu dosen dengan keahlian biomedis dan metabolisme obesitas serta dokter spesialis kedokteran fisik dan rehabilitasi. Pakar menilai relevansi setiap item terhadap kategori SB, NEAT, atau EAT dengan skala 1-4. *Item-level content validity index* (I-CVI) adalah proporsi pakar yang memberi skor 3 atau 4, dan *scale-level content validity index/average* (S-CVI/Ave) adalah rerata seluruh I-CVI. Karena hanya ada dua pakar, item dinyatakan valid jika I-CVI=1,00 (Polit & Beck, 2006; Yusoff, 2019a), dan instrumen dinyatakan valid jika S-CVI/Ave ≥0,90 (Polit & Beck, 2006). Pakar juga menilai ketepatan nilai MET dan kejelasan desain instrumen di luar S-CVI/Ave.

Validitas muka dan reliabilitas diuji pada 20 mahasiswa Program Studi Kedokteran FKIK Universitas Jambi di luar sampel penelitian utama, yaitu 10 mahasiswa dengan status gizi normal

dan 10 mahasiswa *overweight*/obesitas berdasarkan klasifikasi IMT Asia-Pasifik. Responden direkrut secara purposif dari lingkungan pertemanan peneliti. Kriteria inklusi adalah mahasiswa aktif program sarjana berusia 18-25 tahun yang memiliki gawai dan akses internet serta bersedia mengisi *logbook* pada dua minggu berturut-turut. Kriteria eksklusi adalah kehamilan, kondisi medis atau penggunaan obat yang memengaruhi berat badan atau metabolisme energi, riwayat gangguan makan yang didiagnosis secara klinis, dan atlet aktif dengan lima sesi latihan kompetitif atau lebih per minggu. Uji coba dilakukan pada 30 Agustus-1 September 2026 dan 6-8 September 2026, masing-masing pada hari Minggu (hari libur), Senin, dan Selasa, sehingga jarak antarpengukuran adalah tujuh hari. Setelah minggu pertama, responden menilai kejelasan 12 aspek instrumen dengan skala 1-4 melalui Google Form yang disertai pertanyaan terbuka. *Item-level face validity index* (I-FVI) dan *scale-level face validity index/average* (S-FVI/Ave) dihitung dengan cara yang sama seperti CVI, dengan kriteria ≥0,80 (Yusoff, 2019b). Jumlah 20 penilai sudah memenuhi jumlah minimal 10 penilai yang dapat diterima (Yusoff, 2019b). Masukan responden digunakan untuk merevisi instrumen sebelum minggu kedua.

Skor SB adalah total durasi aktivitas SB per hari dalam jam, lalu dirata-ratakan selama tiga hari. Skor NEAT adalah jumlah hasil kali nilai MET dan durasi (menit) aktivitas NEAT per hari, lalu dirata-ratakan selama tiga hari (MET-menit/hari). Skor EAT adalah total menit aktivitas intensitas sedang (3,0-5,9 MET) ditambah dua kali total menit aktivitas intensitas berat (≥6,0 MET) selama tiga hari, lalu dikalikan 7/3 (menit MPA-eq/minggu). Pembobotan ini mengikuti kesetaraan 1 menit aktivitas berat dengan 2 menit aktivitas sedang dalam rekomendasi WHO (Bull et al., 2020; D. Wang et al., 2025). Item EAT di bawah 3,0 MET (yoga dan *workout* mandiri intensitas ringan) tidak dihitung dan tidak dipindahkan ke kategori lain. Isian aktivitas lain dan aktivitas malam dikodekan peneliti ke kode *Compendium* yang sesuai. Aktivitas malam lalu digolongkan sebagai SB atau NEAT berdasarkan nilai MET-nya dan dimasukkan ke hari pelaporannya. Laporan malam tanpa durasi (41 dari 90 laporan yang diisi) tidak dihitung dan tidak diimputasi. Pengiriman ganda pada lima responden diselesaikan dengan memakai entri terakhir. Pada empat dari 120 hari responden, satu dari tiga sesi tidak berisi aktivitas. Sesi tersebut tidak diimputasi, sehingga skor hari itu hanya berasal dari dua sesi yang terisi. Hari dengan total durasi melebihi waktu terjaga (durasi tumpang-tindih) ditemukan pada delapan hari dari empat responden. Data ini tidak dikoreksi pada analisis utama, tetapi keempat responden tersebut dikeluarkan pada analisis sensitivitas.

Reliabilitas *test-retest* dinilai dengan *intraclass correlation coefficient* (ICC) model *two-way mixed*, tipe *absolute agreement*, *single measures* (Koo & Li, 2016), beserta interval kepercayaan 95% (IK95%). Komponen *logbook* dinyatakan reliabel jika ICC ≥0,70, sesuai batas yang umum digunakan (Mokkink, Eekhout, et al., 2023) dan protokol yang ditetapkan sebelum pengambilan data. ICC diinterpretasikan sebagai buruk (<0,50), sedang (0,50-0,75), baik (0,75-

0,90), dan sangat baik (>0,90) dengan mempertimbangkan IK95% (Koo & Li, 2016). Kesalahan pengukuran dinyatakan sebagai *standard error of measurement* (SEM) dan *minimal detectable change* 95%, yaitu MDC95=1,96 ×√2 × SEM (Weir, 2005). SEM ini tidak mencakup perbedaan sistematis antarminggu. Perbedaan antarminggu diuji dengan uji Wilcoxon bertanda (α=0,05) karena sampel kecil dan sebagian selisih tidak terdistribusi normal berdasarkan uji Shapiro-Wilk. Kesesuaian individual digambarkan dengan rerata selisih dan batas kesesuaian 95% Bland-Altman (Bland & Altman, 1986). Analisis sensitivitas dilakukan dengan mengeluarkan satu responden dengan skor EAT ekstrem, mengeluarkan empat responden dengan total durasi harian melebihi waktu terjaga (durasi tumpang-tindih), dan menghitung ICC tanpa periode pukul 21.00 hingga tidur. Analisis dilakukan dengan IBM SPSS Statistics versi 27.

Hasil dan Pembahasan

Karakteristik Responden

Karakteristik responden disajikan pada Tabel 1. Sebagian besar responden adalah perempuan, dan seluruh responden berusia 20-23 tahun. Pada kelompok *overweight*/obesitas, tujuh dari 10 responden tergolong obesitas.

Tabel 1. Karakteristik Responden Uji Coba (n=20)

Karakteristik	Normal (n=10)	*Overweight*/Obesitas (n=10)	Total (n=20)
Jenis kelamin			
Laki-laki	2 (20,0)	1 (10,0)	3 (15,0)
Perempuan	8 (80,0)	9 (90,0)	17 (85,0)
Usia (tahun)	21 (20-22)	21 (20-23)	21 (20-23)
Status gizi (IMT Asia-Pasifik)			
Normal (18,5-22,9 kg/m^2)	10 (100,0)	-	10 (50,0)
Overweight (23,0-24,9 kg/m^2)	-	3 (30,0)	3 (15,0)
Obesitas (≥25,0 kg/m^2)	-	7 (70,0)	7 (35,0)
IMT (kg/m^2)	20,0 (18,6-22,2)	27,9 (24,6-38,6)	23,4 (18,6-38,6)

Data disajikan sebagai n (%) atau median (minimum-maksimum). IMT: Indeks Massa Tubuh.

Validitas Isi dan Validitas Muka

Seluruh 52 item yang dinilai memperoleh I-CVI 1,00 dari kedua pakar, sehingga S-CVI/Ave juga 1,00. Kedua pakar memberi skor 4 untuk relevansi dan ketepatan nilai MET seluruh item, serta untuk seluruh 12 aspek desain dan kejelasan instrumen. Satu-satunya masukan tertulis berasal dari pakar kedokteran fisik dan rehabilitasi, yaitu memisahkan item HIIT dan *calisthenics* keras, sehingga *logbook* versi akhir berisi 53 item. Pada uji validitas muka, S-FVI/Ave sebesar 0,97, dengan I-FVI setiap aspek 0,90-1,00. Nilai terendah (0,90) terdapat pada aspek cara mencatat aktivitas naik-turun tangga, sehingga petunjuknya diperjelas dan diberi contoh sebelum minggu kedua. Pada pertanyaan terbuka, beberapa responden kesulitan menghitung

total menit secara manual, kurang familier dengan aktivitas EAT, dan belum memahami beberapa kode atau istilah, meskipun contoh yang disediakan dinilai membantu. Kendala teknis yang dilaporkan adalah tampilan formulir yang kurang optimal di layar ponsel sehingga sempat terjadi salah isi, gangguan jaringan, dan *error* sesaat. Kendala ini dianggap masalah teknis dan tidak dijadikan dasar revisi isi instrumen.

Hasil ini menunjukkan bahwa daftar aktivitas dinilai relevan oleh pakar (Polit & Beck, 2006; Yusoff, 2019a) dan petunjuk *logbook* dapat dipahami oleh responden (Yusoff, 2019b). Namun, validitas muka hanya menilai kejelasan instrumen. Hasil uji reliabilitas berikut menunjukkan bahwa *logbook* yang dinilai jelas tetap dapat menghasilkan pelaporan durasi yang tidak konsisten, kemungkinan karena responden kesulitan menghitung total menit dan salah isi di layar ponsel yang kecil.

Reliabilitas *Test-Retest*

Hasil uji reliabilitas disajikan pada Tabel 2. ICC EAT berada pada kategori baik, ICC NEAT pada kategori sedang, dan ICC SB pada kategori buruk. ICC NEAT memenuhi kriteria ≥0,70, tetapi batas bawah IK95%-nya masih tergolong buruk, sedangkan ICC SB belum memenuhi kriteria. Skor NEAT minggu kedua lebih rendah secara bermakna, sedangkan skor SB dan EAT tidak berbeda bermakna. Batas kesesuaian Bland-Altman ketiga variabel cukup lebar, sehingga kesesuaian individual masih terbatas. Distribusi EAT sangat miring ke kanan. Sebanyak 10 responden memiliki skor EAT 0 pada kedua minggu (11 responden pada minggu pertama dan 14 responden pada minggu kedua), sehingga median EAT pada kedua minggu adalah 0 menit MPA-eq/minggu. Selisih antarminggu terdistribusi normal untuk SB dan NEAT (uji Shapiro-Wilk, p=0,312 dan p=0,401), tetapi tidak untuk EAT (p=0,001), sehingga batas kesesuaian Bland-Altman EAT kurang dapat diandalkan.

Tabel 2. Hasil Uji Reliabilitas *Test-Retest* (n=20)

Variabel	Minggu 1 (rerata ± SD)	Minggu 2 (rerata ± SD)	Selisih rerata (BK95%)	Nilai p	ICC (IK95%)	SEM	MDC95
SB (jam/hari)	8,09 ± 3,33	8,37 ± 3,01	0,28 (-6,06;6,62)	0,502	0,491 (0,065-0,764)	2,29	6,34
NEAT (MET-menit/hari)	465,7 ± 180,4	376,4 ± 263,7	-89,2 (-391,3;212,8)	0,019	0,720 (0,363-0,884)	109,0	302,0
EAT (menit MPA-eq/minggu)	187,8 ± 381,7	158,1 ± 454,6	-29,8 (-405,8;346,3)	0,406	0,898 (0,763-0,958)	135,7	376,1

Selisih = minggu 2-minggu 1; BK95%: batas kesesuaian 95% Bland-Altman; nilai p: uji Wilcoxon bertanda; ICC: model *two-way mixed, absolute agreement, single measures*; IK95%: interval kepercayaan 95%; SEM: *standard error of measurement*; MDC95: *minimal detectable change* 95%. Satuan SEM dan MDC95 mengikuti satuan variabel. Skor SB dan NEAT mencakup periode pukul 21.00 hingga tidur; skor EAT tidak mencakup item <3,0 MET. Selisih rerata, SEM, dan MDC95 dihitung dari nilai yang tidak dibulatkan.

Skor NEAT minggu kedua sekitar 19% lebih rendah, dan penurunan terjadi pada 14 dari 20 responden. Nilai MDC95 NEAT setara dengan sekitar 72% dari rerata skor NEAT kedua minggu. Penurunan terbesar terjadi pada hari Senin dengan rerata selisih -261,0 MET-menit, dibandingkan dengan -52,2 MET-menit pada hari Minggu dan +45,5 MET-menit pada hari Selasa. Pada Senin minggu pertama, 16 responden mencatat kuliah tatap muka pada item kuliah aktif dengan rerata 167,5 menit per responden. Pada Senin minggu kedua, kuliah berlangsung daring dan hanya tiga responden yang mencatatnya sebagai isian SB lainnya. Total durasi yang tercatat pada hari Senin juga turun dari 11,6 menjadi 9,1 jam per responden (uji Wilcoxon bertanda, p=0,015). Durasi naik tangga, satu-satunya item yang petunjuknya direvisi, tidak berbeda antara kedua minggu (3,8 dan 4,0 menit/hari, p=0,856). Hari libur minggu kedua (6 September 2026) bertepatan dengan musyawarah besar organisasi kemahasiswaan yang diikuti 12 dari 20 responden. Pada hari itu, SB delapan responden naik lebih dari 2 jam dibandingkan hari libur minggu pertama, sedangkan SB tujuh responden turun lebih dari 2 jam.

Hasil analisis sensitivitas disajikan pada Tabel 3. Setelah satu responden dengan skor EAT ekstrem dikeluarkan, ICC EAT turun ke kategori buruk dan tidak lagi berbeda bermakna dari nol (p=0,063). ICC SB naik ke kategori sedang pada ketiga analisis sensitivitas, sedangkan ICC NEAT hampir tidak berubah.

Tabel 3. Analisis Sensitivitas ICC

Variabel	**Seluruh responden (n=20)**	**Tanpa responden ekstrem (n=19)**	**Tanpa responden dengan durasi tumpang-tindih (n=16)**	**Tanpa periode pukul 21.00 hingga tidur (n=20)**
SB	0,491 (0,065-0,764)	0,520 (0,108-0,782)	0,581 (0,156-0,828)	0,564 (0,163-0,803)
NEAT	0,720 (0,363-0,884)	0,719 (0,354-0,886)	0,730 (0,058-0,918)	0,706 (0,332-0,879)
EAT	0,898 (0,763-0,958)	0,346 (-0,094 sampai 0,680)	0,323 (-0,167 sampai 0,693)	0,898 (0,763-0,958)

Nilai dalam tabel adalah ICC (IK95%) dengan model *two-way mixed, absolute agreement, single measures*. Skor EAT tidak mencakup periode pukul 21.00 hingga tidur, sehingga ICC-nya sama pada kolom terakhir.

Reliabilitas SB pada penelitian ini tergolong buruk, tetapi setara dengan reliabilitas SB pada *Global Physical Activity Questionnaire* (GPAQ) dan IPAQ versi pendek (IPAQ-SF) pada mahasiswa di Tiongkok dengan interval delapan hari, yaitu ICC 0,47 dan 0,52 (Gao et al., 2022). Tinjauan sistematis juga melaporkan reliabilitas SB yang sangat bervariasi pada alat swalapor perilaku gerak 24 jam, yaitu 0,33-0,91 (Šuc et al., 2024). Nilai yang lebih tinggi dilaporkan pada *time-use diary* 24 jam pada pekerja kantoran di Thailand, yaitu ICC SB 0,70 pada seluruh sampel dan 0,83 pada responden dengan aktivitas serupa (Liangruenrom et al., 2025). Kuesioner SB waktu luang mahasiswa di Kolombia bahkan menunjukkan ICC 0,93 (Atencio-Osorio et al., 2021). Namun, perbandingan ini perlu dibaca hati-hati karena interval, populasi, dan sebagian model ICC setiap studi berbeda.

Karena *logbook* mencatat perilaku pada tiga hari tertentu, ICC antarminggu dipengaruhi

oleh kesalahan pengukuran dan juga oleh perubahan perilaku yang nyata. Perbedaan aktivitas antarperiode diketahui dapat menurunkan reliabilitas SB (Liangruenrom et al., 2025), dan frekuensi duduk lama (≥2 jam) pada mahasiswa dilaporkan berubah secara tidak linear dari minggu ke minggu (Garn & Simonton, 2023). Pada studi dengan akselerometer pun, ICC waktu sedentari antarminggu hanya 0,52 sebelum dikoreksi dengan waktu pemakaian alat dan 0,71 setelah dikoreksi (Aadland & Ylvisåker, 2015). Reliabilitas 0,80 pada waktu sedentari juga baru tercapai dengan sekitar delapan hari pengamatan (Hilden et al., 2023), sedangkan penelitian ini hanya menggunakan tiga hari. Tiga hari tetap dipilih karena rerata aktivitas fisik kelompok selama tiga hari (satu hari libur dan dua hari kerja) tidak berbeda bermakna dengan tujuh hari pada mahasiswa, walaupun waktu sedentari tidak dinilai (Han & Dinger, 2009). Hari libur disertakan karena pola aktivitasnya sedikit berbeda dari hari kerja (Aadland & Ylvisåker, 2015), dan pencatatan yang lebih lama dikhawatirkan terlalu membebani responden.

Kondisi kedua minggu juga tidak sepenuhnya sama. Pada hari libur minggu kedua yang bertepatan dengan musyawarah besar, perubahan SB terjadi ke dua arah, sehingga rerata kelompok tidak banyak berubah tetapi konsistensi antarindividu menurun. Kuliah daring pada Senin minggu kedua yang sebagian besar tidak tercatat kemungkinan juga membuat sebagian SB tidak terlaporkan. ICC SB naik ketika periode pukul 21.00 hingga tidur tidak dimasukkan, kemungkinan karena periode ini hanya dilaporkan melalui *recall* satu aktivitas dominan dan hampir separuh laporannya tanpa durasi. ICC SB juga naik ke kategori sedang setelah responden ekstrem atau responden dengan durasi tumpang-tindih dikeluarkan, sehingga pelaporan yang tumpang-tindih kemungkinan ikut menurunkan reliabilitas SB. Meskipun begitu, ICC SB masih berada dalam rentang 0,44-0,91 yang dilaporkan untuk metode subjektif SB, walaupun di batas bawah (Bakker et al., 2020).

ICC NEAT pada penelitian ini memenuhi kriteria. Belum ada instrumen lain yang mengukur NEAT secara terpisah, tetapi nilai ini lebih tinggi daripada ICC aktivitas intensitas sedang hingga berat pada domain kerja, transportasi, dan rumah tangga pada GPAQ dan IPAQ versi panjang, yaitu 0,08-0,32 (Gao et al., 2022). Nilai ini juga berada dalam rentang reliabilitas aktivitas fisik ringan, yaitu 0,55-0,94 (Šuc et al., 2024), walaupun definisinya tidak sama persis. Meskipun begitu, skor NEAT minggu kedua lebih rendah secara bermakna. Karena ICC *absolute agreement* memperhitungkan perbedaan sistematis, penurunan ini membuat ICC lebih rendah (Koo & Li, 2016; Mokkink, Eekhout, et al., 2023). Penurunan ini sebagian berasal dari kuliah Senin yang berpindah ke daring, sedangkan *logbook* belum memiliki item kuliah daring, sehingga sebagian besar responden tidak mencatatnya. Penurunan total durasi yang tercatat juga menunjukkan pelaporan yang kurang lengkap. Hal ini memang menjadi tantangan umum dalam EMA, dengan rerata kepatuhan sekitar 81% pada studi EMA perilaku gerak remaja dan dewasa muda (S. Wang et al., 2025). MDC95 NEAT yang besar dibandingkan rerata skornya juga

menunjukkan bahwa *logbook* belum cukup presisi untuk menilai perubahan NEAT individu.

ICC EAT pada penelitian ini termasuk tinggi. Sebagai pembanding, ICC aktivitas intensitas sedang hingga berat adalah 0,66-0,86 untuk waktu luang pada GPAQ dan IPAQ versi panjang (Gao et al., 2022) dan 0,40 pada *time-use diary* di Thailand (Liangruenrom et al., 2025). Namun, nilai ICC bergantung pada besarnya variasi antarsubjek (Koo & Li, 2016; Weir, 2005), sehingga satu responden dengan skor EAT yang jauh lebih tinggi dapat menaikkan ICC walaupun kesesuaian responden lain rendah. Hal ini terlihat dari turunnya ICC EAT ke kategori buruk setelah responden tersebut dikeluarkan. Setengah responden juga memiliki skor EAT 0 pada kedua minggu. Nilai nol ini menaikkan kesesuaian secara statistik, tetapi tidak menggambarkan konsistensi pada mahasiswa yang berolahraga. Olahraga juga tidak dilakukan setiap hari, sehingga skor EAT dari tiga hari dengan pengali 7/3 sangat bergantung pada ada atau tidaknya sesi olahraga pada hari yang tercatat. Untuk aktivitas intensitas sedang hingga berat, dibutuhkan sekitar sepuluh hari pengamatan untuk mencapai reliabilitas 0,80 (Hilden et al., 2023).

Secara umum, *logbook* ini memiliki validitas isi dan validitas muka yang sangat baik, tetapi reliabilitasnya berbeda pada setiap komponen, sehingga hasil ini baru menjadi bukti awal. Untuk saat ini, skor *logbook* lebih tepat digunakan untuk membandingkan kelompok daripada menilai individu. Hal ini sejalan dengan temuan bahwa alat swalapor perilaku gerak sudah memadai untuk studi epidemiologis dan surveilans populasi, tetapi kecukupannya untuk penilaian individual belum banyak diketahui (Liangruenrom et al., 2025; Šuc et al., 2024). Reliabilitas SB yang rendah juga dapat melemahkan estimasi hubungan SB dengan luaran kesehatan pada penelitian utama. Perbaikan yang disarankan adalah menambahkan item kuliah daring, validasi otomatis di KoBoToolbox agar total durasi per sesi tidak melebihi rentang sesi, format pencatatan periode malam yang sama dengan sesi lain, serta periode pencatatan EAT yang lebih panjang atau pertanyaan tentang kebiasaan olahraga mingguan. Item kuliah daring (Zoom/Google Meet) sudah ditambahkan ke kategori SB sebelum penelitian utama dengan kode 11582 (*sitting, computer work,* 1,3 MET) dan telah dikonfirmasi oleh pakar kedokteran fisik dan rehabilitasi. Dengan klasifikasi ini, kuliah daring tergolong SB, sedangkan kuliah tatap muka yang aktif (1,8 MET) tetap tergolong NEAT.

Penelitian ini memiliki beberapa keterbatasan. Pertama, jumlah responden masih di bawah patokan minimal 30 sampel untuk studi reliabilitas (Koo & Li, 2016) dan perkiraan 50-100 responden untuk desain dua kali pengukuran, sehingga IK95% ICC lebar (Mokkink, de Vet, et al., 2023). Kedua, validitas isi hanya dinilai oleh dua pakar yang memiliki hubungan akademik dengan peneliti, padahal dianjurkan minimal tiga pakar (Polit & Beck, 2006; Roebianto et al., 2023), sehingga penilaiannya mungkin tidak sepenuhnya independen, apalagi seluruh item mendapat skor maksimal. Dua item EAT hasil pemecahan belum dinilai melalui CVI dan kode MET-nya hanya disetujui oleh satu pakar. Item kuliah daring juga hanya dikonfirmasi oleh satu

pakar dan belum diuji reliabilitasnya. Ketiga, validitas kriteria terhadap metode objektif, seperti akselerometer atau *doubly labeled water*, belum dinilai. Keempat, desain *test-retest* dalam kondisi sehari-hari tidak dapat memisahkan kesalahan pengukuran dari perubahan perilaku, apalagi kondisi kedua minggu tidak sepenuhnya sama dan petunjuk item naik tangga direvisi di antara kedua minggu. Tiga hari pencatatan juga masih kurang untuk penilaian individual, terutama untuk EAT. Kelima, periode pukul 21.00 hingga tidur hanya dicatat melalui *recall* satu aktivitas dominan, hampir separuh laporannya tanpa durasi, dan periode malam hari ketiga tidak tercakup, sehingga SB dan NEAT malam hari kemungkinan lebih rendah dari yang sebenarnya. Keenam, data hari dengan durasi tumpang-tindih tetap dipertahankan dalam analisis utama. Ketujuh, nilai MET untuk salat ditentukan berdasarkan kode aktivitas terdekat pada *Compendium*. Kedelapan, responden berasal dari satu program studi dan direkrut dari lingkungan pertemanan peneliti, sehingga hasilnya belum tentu berlaku pada mahasiswa lain.

Kesimpulan

Logbook aktivitas fisik digital berbasis EMA ini memiliki validitas isi dan validitas muka yang sangat baik, tetapi reliabilitas *test-retest*-nya berbeda pada setiap komponen. ICC NEAT hanya sedikit di atas kriteria, ICC SB tergolong buruk, dan ICC EAT tergolong baik tetapi tidak stabil karena sangat dipengaruhi oleh satu responden. Sebagian perbedaan antarminggu kemungkinan berkaitan dengan perubahan kondisi perkuliahan, kegiatan kemahasiswaan, dan belum adanya item kuliah daring. Instrumen ini dapat digunakan untuk analisis tingkat kelompok, tetapi skor SB perlu dibaca dengan hati-hati. Penelitian selanjutnya disarankan menyempurnakan item dan cara pencatatan *logbook*, serta menguji validitas kriteria terhadap metode objektif pada sampel yang lebih besar.

Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Unit Teknologi Informasi FKIK Universitas Jambi atas dukungan validasi teknis platform, serta kepada seluruh mahasiswa yang berpartisipasi sebagai responden.

Referensi

Aadland, E., & Ylvisåker, E. (2015). Reliability of objectively measured sedentary time and physical activity in adults. *PLOS ONE*, *10*(7), e0133296. https://doi.org/10.1371/journal.pone.0133296

Akbar, M. S., Jajat, J., Sultoni, K., Ruhayati, Y., Suherman, A., & Nuryanti, W. D. (2025). Validity and reliability of the Indonesian version of the International Physical Activity Questionnaire among students: A confirmatory analysis using the Actigraph GT3X+ accelerometer. *Gladi Jurnal Ilmu Keolahragaan*, *16*(1), 25–33. https://doi.org/10.21009/GJIK.161.03

Amornsriwatanakul, A., Rahman, H. A., Wattanapisit, A., Nurmala, I., de la Cruz, M. H. T. O., Car, J., & Chia, M. (2022). University students' overall and domain-specific physical activity during COVID-19: A cross-sectional study in seven ASEAN countries. *Heliyon*, *8*(12), e12466.

https://doi.org/10.1016/j.heliyon.2022.e12466

Atencio-Osorio, M. A., Carrillo-Arango, H. A., Correa-Rodríguez, M., Rivera, D., Castro-Piñero, J., & Ramírez-Vélez, R. (2021). Youth Leisure-Time Sedentary Behavior Questionnaire (YLSBQ): Reliability and validity in Colombian university students. *International Journal of Environmental Research and Public Health*, *18*(15), 7895. https://doi.org/10.3390/ijerph18157895

Bakker, E. A., Hartman, Y. A. W., Hopman, M. T. E., Hopkins, N. D., Graves, L. E. F., Dunstan, D. W., Healy, G. N., Eijsvogels, T. M. H., & Thijssen, D. H. J. (2020). Validity and reliability of subjective methods to assess sedentary behaviour in adults: A systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 75. https://doi.org/10.1186/s12966-020-00972-1

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 327*(8476), 307-310. https://doi.org/10.1016/S0140-6736(86)90837-8

Bonilla, D. A., Peralta-Alzate, J. O., Bonilla-Henao, J. A., Cannataro, R., Cardozo, L. A., Vargas-Molina, S., Stout, J. R., Kreider, R. B., & Petro, J. L. (2023). Insights into non-exercise physical activity on control of body mass: A review with practical recommendations. *Journal of Functional Morphology and Kinesiology*, *8*(2), 44. https://doi.org/10.3390/jfmk8020044

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J. P., Chastin, S., Chou, R., Dempsey, P. C., Dipietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., … Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine*, *54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health Reports*, *100*(2), 126–131.

Chung, N., Park, M.-Y., Kim, J., Park, H.-Y., Hwang, H., Lee, C.-H., Han, J.-S., So, J., Park, J., & Lim, K. (2018). Non-exercise activity thermogenesis (NEAT): A component of total daily energy expenditure. *Journal of Exercise Nutrition & Biochemistry*, *22*(2), 23–30. https://doi.org/10.20463/jenb.2018.0013

Gao, H., Li, X., Zi, Y., Mu, X., Fu, M., Mo, T., & Yu, K. (2022). Reliability and validity of common subjective instruments in assessing physical activity and sedentary behaviour in Chinese college students. *International Journal of Environmental Research and Public Health*, *19*(14), 8379. https://doi.org/10.3390/ijerph19148379

Garn, A. C., & Simonton, K. L. (2023). Prolonged sitting in university students: An intra-individual study exploring physical activity value as a deterrent. *International Journal of Environmental Research and Public Health*, *20*(3), 1891. https://doi.org/10.3390/ijerph20031891

Hamasaki, H., Yanai, H., Kakei, M., Noda, M., & Ezaki, O. (2014). The validity of the non-exercise activity thermogenesis questionnaire evaluated by objectively measured daily physical activity by the triaxial accelerometer. *BMC Sports Science, Medicine and Rehabilitation*, *6*, 27. https://doi.org/10.1186/2052-1847-6-27

Han, J. L., & Dinger, M. K. (2009). Validity of a self-administered 3-day physical activity recall in young adults. *American Journal of Health Education*, *40*(1), 5–13. https://doi.org/10.1080/19325037.2009.10599073

Hartson, K. R., Huntington-Moskos, L., Sears, C. G., Genova, G., Mathis, C., Ford, W., & Rhodes, R. E. (2023). Use of electronic ecological momentary assessment methodologies in physical activity, sedentary behavior, and sleep research in young adults: Systematic review. *Journal of Medical Internet Research*, *25*, e46783. https://doi.org/10.2196/46783

Herrmann, S. D., Willis, E. A., Ainsworth, B. E., Barreira, T. V., Hastert, M., Kracht, C. L., Schuna, J. M., Jr., Cai, Z., Quan, M., Tudor-Locke, C., Whitt-Glover, M. C., & Jacobs, D. R., Jr. (2024). 2024 Adult Compendium of Physical Activities: A third update of the energy costs of human activities. *Journal of Sport and Health Science*, *13*(1), 6–12. https://doi.org/10.1016/j.jshs.2023.10.010

Hilden, P., Schwartz, J. E., Pascual, C., Diaz, K. M., & Goldsmith, J. (2023). How many days are needed? Measurement reliability of wearable device data to assess physical activity. *PLOS ONE*, *18*(2), e0282162. https://doi.org/10.1371/journal.pone.0282162

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163.

https://doi.org/10.1016/j.jcm.2016.02.012

Liangruenrom, N., Suttikasem, K., Widyastari, D. A., Potharin, D., & Katewongsa, P. (2025). Reliability and validity of time-use surveys in assessing 24-hour movement behaviors in adults. *Journal of Exercise Science and Fitness*, *23*(2), 133–140. https://doi.org/10.1016/j.jesf.2025.03.003

Meh, K., Jurak, G., Sorić, M., Rocha, P., & Sember, V. (2021). Validity and reliability of IPAQ-SF and GPAQ for assessing sedentary behaviour in adults in the European Union: A systematic review and meta-analysis. *International Journal of Environmental Research and Public Health*, *18*(9), 4602. https://doi.org/10.3390/ijerph18094602

Mokkink, L. B., de Vet, H., Diemeer, S., & Eekhout, I. (2023). Sample size recommendations for studies on reliability and measurement error: An online application based on simulation studies. *Health Services and Outcomes Research Methodology*, *23*(3), 241–265. https://doi.org/10.1007/s10742-022-00293-9

Mokkink, L. B., Eekhout, I., Boers, M., van der Vleuten, C. P. M., & de Vet, H. C. W. (2023). Studies on reliability and measurement error of measurements in medicine - From design to statistics explained for medical researchers. *Patient Related Outcome Measures*, *14*, 193–212. https://doi.org/10.2147/PROM.S398886

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147

Putri, R. F. K., Hanifah, F. N. N., Devina, K. E., & Kusumawati, D. (2026). The relationship between skipping meals, sedentary behavior, and sleep duration with the incidence of overweight in final-year university students. *Jurnal Riset Gizi*, *Vol. 14 (1)*, 22–29. https://ejournal.poltekkes-smg.ac.id/ojs/index.php/jrg/article/view/13283/8962

Rizzato, A., Marcolin, G., & Paoli, A. (2022). Non-exercise activity thermogenesis in the workplace: The office is on fire. *Frontiers in Public Health*, *10*, 1024856. https://doi.org/10.3389/fpubh.2022.1024856

Roebianto, A., Savitri, S. I., Sriyanto, A., Syaiful, I., & Mubarokah, L. (2023). Content validity: Definition and procedure of content validation in psychological research. *TPM-Testing, Psychometrics, Methodology in Applied Psychology*, *30*(1), 5–18. https://doi.org/10.4473/TPM30.1.1

Setyaningsih, A., & Fatqur, A. (2025). Faktor risiko gizi lebih pada dewasa muda: Konsumsi minuman manis dan perilaku sedentari. *Jurnal Riset Gizi*, *13*(1), 59–64. https://doi.org/10.31983/jrg.v13i1.12941

Šuc, A., Einfalt, L., Šarabon, N., & Kastelic, K. (2024). Validity and reliability of self-reported methods for assessment of 24-h movement behaviours: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, *21*(1), 83. https://doi.org/10.1186/s12966-024-01632-4

Telleria-Aramburu, N., & Arroyo-Izaga, M. (2022). Risk factors of overweight/obesity-related lifestyles in university students: Results from the EHU12/24 study. *British Journal of Nutrition*, *127*(6), 914–926. https://doi.org/10.1017/S0007114521001483

Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., Chastin, S. F. M., Altenburg, T. M., & Chinapaw, M. J. M. (2017). Sedentary Behavior Research Network (SBRN) - Terminology Consensus Project process and outcome. *International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 75. https://doi.org/10.1186/s12966-017-0525-8

von Loeffelholz, C., & Birkenfeld, A. L. (2022). Non-exercise activity thermogenesis in human energy homeostasis. In K. R. Feingold et al. (Eds.), Endotext. MDText.com, Inc. https://www.ncbi.nlm.nih.gov/books/NBK279077/

Wang, D., Wu, L., Yan, L., Yang, H., Huang, X., Wang, Z., & Guan, Y. (2025). Associations of the volume and proportion of vigorous-intensity physical activity with all-cause, cardiovascular, and cancer mortality: A systematic review and meta-analysis. *PeerJ*, *13*, e19538. https://doi.org/10.7717/peerj.19538

Wang, S., Yang, C.-H., Brown, D., Cheng, A., & Kwan, M. Y. W. (2025). Participant compliance with ecological momentary assessment in movement behavior research among adolescents and emerging adults: Systematic review. *JMIR mHealth and uHealth*, *13*, e52887. https://doi.org/10.2196/52887

Weir, J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the

SEM. *Journal of Strength and Conditioning Research*, *19*(1), 231–240. https://doi.org/10.1519/15184.1

World Health Organization. (2026, October 7). Obesity and overweight. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

Yulinda, I., Ayudia, E. I., Kusdiyah, E., Harahap, H., & Perkasa, T. A. B. (2026). Faktor risiko overweight dan obesitas pada mahasiswa kedokteran Universitas Jambi angkatan 2022-2024. *Jurnal Medika Malahayati*, *10*(2), 270–281. https://doi.org/10.33024/jmm.v10i2.24451

Yuniarti, T., Musta'in, Adriani, R. B., Widiyanto, A., Atmojo, J. T., & Putri, S. I. (2024). The correlation between physical activity, sedentary behavior, and body mass index among college students in Surakarta, Indonesia: A cross-sectional study. *Asian Journal of Social Health and Behavior*, *7*(3), 134–139. https://doi.org/10.4103/shb.shb_18_24

Yusoff, M. S. B. (2019a). ABC of content validation and content validity index calculation. *Education in Medicine Journal*, *11*(2), 49–54. https://doi.org/10.21315/eimj2019.11.2.6

Yusoff, M. S. B. (2019b). ABC of response process validation and face validity index calculation. *Education in Medicine Journal*, *11*(3), 55–61. https://doi.org/10.21315/eimj2019.11.3.6