From Stability to Complexity: A Systematic Review of Long-term Divergence Exponents in Nonlinear Gait Analysis

Jeremy Torrent[1¶], Roxane Coquoz[1¶], Philippe Terrier[1*]

1 Haute-Ecole Arc Santé, HES-SO University of Applied Sciences and Arts Western Switzerland, Neuchâtel, Switzerland

¶ These authors contributed equally to this work.

* Corresponding Author:

Philippe Terrier
Haute-Ecole Arc Santé
Espace de l'Europe 11, CH-2000 Neuchâtel, Switzerland
Philippe.terrier@he-arc.ch

Keywords: Human locomotion; Walking; Gait analysis; Gait variability; Gait automaticity; Gait complexity; Nonlinear dynamics; Lyapunov exponent

Abstract

Background

Divergence exponents (DE), or maximum Lyapunov exponents, computed from stride-to-stride fluctuations have traditionally been interpreted as measures of gait stability. However, evidence suggests this measure may also reflect gait complexity and automaticity. This systematic review evaluated empirical support for reinterpreting long-term DE as a complexity measure.

Methods

We systematically searched Web of Science databases through January 2026 for studies applying Rosenstein's algorithm to human gait. We systematically extracted experimental conditions, and participant characteristics from each study. Study quality was assessed using a framework evaluating analytical rigor, outcome reporting, and sample size adequacy. We conducted a meta-analysis examining correlations between long-term DEs and detrended fluctuation analysis (DFA) scaling exponents, and synthesized evidence from perturbation (environmental disturbances), cueing (external rhythmic stimuli), and between-subject (clinical vs control) studies.

Results

Sixty-two studies published between 2000 and 2026 met inclusion criteria, with 44% achieving high overall quality scores. Meta-analysis from six datasets (209 participants) revealed a positive correlation between long-term DE and DFA scaling exponents ($r=0.64$, 95% CI 0.34 to 0.82; $I^2=82\%$). Perturbation studies consistently showed increased short-term DE (lower local stability) while simultaneously decreasing long-term DE by up to 51%. External auditory and visual cueing interventions induced long-term DE decreases (up to -86%) while minimally affecting short-term DEs. Between-subject comparisons revealed heterogeneous patterns, with clinical populations exhibiting both increases and decreases in long-term DE depending on pathology.

Conclusions

Converging meta-analytic and experimental evidence supports reinterpreting long-term DE as the Attractor Complexity Index—a measure of gait complexity and automaticity rather than stability. Reduced long-term DE during controlled gait conditions reflects suppression of low-frequency variations in gait variability, which constrains phase space exploration and accelerates divergence curve saturation. The measure's sensitivity to prefrontal cortex engagement and attentional demands reveals that long-term DE captures the degree of gait

automaticity, with lower values indicating a shift from automatic subcortical control to executive function-mediated regulation. This attention-dependent control reorganization explains patterns observed in aging, clinical populations, and experimental perturbations. This paradigm shift establishes long-term DE as a complementary biomarker for motor-cognitive aspects of gait control, with implications for fall risk assessment, disease monitoring, and rehabilitation evaluation.

Introduction

Walking is a complex motor activity that plays an important role in daily functioning [1]. Gait—the pattern of limb movement—can serve as a significant biomarker, providing insights into an individual's physical functioning [2,3], cognitive abilities [4], and level of independence [5]. Slower speeds and deviations from normal gait patterns have been shown to correlate with adverse outcomes, including an increased risk of falls [6,7], cognitive decline [7,8], and mortality [9]. For this reason, gait analysis is an essential component of clinical evaluation across medical specialties [10,11].

Nonlinear gait analysis is a methodological approach that examines the seemingly unpredictable dynamics of human gait using mathematical and statistical tools from nonlinear systems theory [12]. Nonlinear dynamical systems are characterized by outputs that are not proportional to changes in inputs, where the behavior cannot be described by simple linear differential equations. In the context of physiological signals, nonlinearity reflects how small changes in one component can lead to disproportionate or emergent effects throughout the system. These complex interactions arise from the integration of multiple regulatory mechanisms—such as the interplay between the autonomic nervous system and neuroendocrine factors in heart rate regulation—forming a dynamic network whose behavior cannot be fully understood by analyzing individual components in isolation [13,14]. Nonlinear analysis techniques have become an integral part of gait analysis, providing advanced methods such as maximum Lyapunov (or divergence) exponents for assessing resilience to perturbations [15], fractal dynamics for investigating the complexity and statistical persistence of time series [16], and entropy measures for quantifying signal irregularity [17]. These methods facilitate the detection of gait disorders [18–20], and support the development of innovative rehabilitation strategies [21]. By going beyond traditional kinematic and kinetic parameters, nonlinear approaches allow for a more comprehensive assessment of gait variability [22].

Nonlinear gait analysis also provides a window into the underlying neural control mechanisms [23], including the involvement of executive function in gait [4]. Attention, as a component of executive function, refers to the ability to allocate cognitive resources among tasks, including gait, that are performed concurrently [4,24]. Gait automaticity reflects the extent to which executive function contributes to gait control, with highly automatic gait requiring little attentional resources and less automatic gait requiring increased executive involvement [25]. One particularly powerful approach for investigating the interplay between executive function and gait is the use of dual-task paradigms. In these experimental setups,

participants are asked to perform a cognitive task while walking, which allows researchers to assess how the allocation of attentional resources affects gait parameters [4]. Nonlinear measures applied to dual-task walking data can reveal how the complexity and stability of gait are modulated by concurrent cognitive demands, providing valuable information about the adaptive capacity of the motor control system and the degree of gait automaticity [26,27]. This approach has proven useful in understanding age-related changes in gait control [28] and in identifying early markers of cognitive decline [29] and fall risk [30] in various clinical populations.

Beyond dual-task paradigms, various factors—including sensory inputs, environmental challenges, and individual physiological and psychological states—can influence the extent of executive function involvement during walking, and nonlinear gait analysis provides valuable tools for exploring these dynamics. For example, the influence of auditory cues on gait automaticity can be observed through rhythmic auditory stimulation, where external rhythms entrain motor rhythms, potentially bypassing impaired internal gait control mechanisms in conditions such as Parkinson's disease [31,32]. Exposing older adults to visual perturbations significantly reduces the complexity of their gait, a response not observed in younger adults. This may be indicative of an increased reliance on visual feedback and a decrease in gait automaticity with age [33]. Environmental constraints, such as treadmill walking, can alter motor control strategies. Research has shown that stride-speed fluctuations exhibit statistical anti-persistence during treadmill walking, indicating tight control over this gait parameter. This tight control is further enhanced when maneuverability is restricted, suggesting that the nervous system actively minimizes speed deviations [34]. These findings are consistent with the idea that the need for continuous speed adjustments on a treadmill may increase the demands on executive function to maintain a steady gait [35]. Individual physiological and psychological states, such as pain, attentional focus, or anxiety, may also influence gait automaticity. For example, research suggests that individuals with chronic low back pain exhibit altered gait complexity, possibly due to attentional strategies focused on managing pain [36]. In addition, fear of falling is associated with reduced local dynamic stability of gait [37], which may be explained by an increased reliance on executive functions to maintain balance and adapt to perceived threats during walking.

Among nonlinear measures, the maximum Lyapunov exponent [divergence exponent (DE)] method stands out as a prominent measure for assessing dynamic stability and fall risk [20,38,39]. This prominence is due to its ability to quantify local dynamic stability—a system's sensitivity to small perturbations [15,40,41]. Modeling studies using dynamic walker systems have validated its predictive validity, demonstrating strong correlations between

short-term DEs and simulated fall probabilities [42,43]. These computational validations, combined with clinical observations of elevated DE in fall-prone elderly, position it as an important tool for assessing gait stability and fall risk in both research and clinical applications [20]. However, recent systematic reviews have shown heterogeneity in the implementation of the DE method, with different parameterizations and measurement systems across studies [20,44], highlighting the need for further research to standardize methodologies and facilitate cross-study comparisons.

A persistent methodological challenge in DE-based gait analysis arises from the differential responsiveness of short- and long-term DE computed by Rosenstein's algorithm. Mathematically, Rosenstein's method embeds the time series in a state space and identifies for each point a nearest neighbor such that their initial separation $d(0)$ evolves with time t as $d(t) = d(0)e^{\lambda t}$, where λ is the DE characterizing the intensity of the exponential divergence, which is also known as the "butterfly effect" [45]; taking the natural logarithm gives $ln(d(t)) = ln(d(0)) + \lambda\, t$, and hence $\lambda \approx \frac{\ln(d(t))}{\Delta t}$ [15,40]. The characteristic logarithmic divergence curve is computed to assess the average progressive separation. Short-term DE is computed as the average rate of divergence by linear regression of the initial segment of the curve (0-1 step or stride), capturing local dynamic stability through immediate corrective mechanisms. Conversely, long-term DE is derived from later intervals (4-10 strides) and reflects long-term system behavior. While elevated short-term DE consistently correlates with fall risk and other variability metrics [46], long-term DE exhibits paradoxical responses to experimental manipulations, decreasing during cognitively demanding walking [33,47] and showing weak, or even opposite, associations with stability measures [46,48].

Therefore, while long-term DE was initially designed to measure gait stability, it has been reinterpreted as capturing distinct aspects of gait control compared to short-term DE. Several studies have demonstrated strong associations between long-term DE and the correlation structure of stride intervals, as quantified by scaling exponents that characterize stride-to-stride fluctuation complexity [40,49,50]. This relationship is particularly evident in experimental conditions that modify gait automaticity. Metronome walking, which requires conscious sensorimotor synchronization and increased attentional demands, substantially decreases both long-term DE and gait complexity while leaving short-term DE unchanged [47,51]. Furthermore, it has been shown that gait acceleration signals with identical shapes, but modified stride interval complexity exhibit different long-term DE values [52]. These findings have led to its reconceptualization as the "Attractor Complexity Index" (ACI), reflecting its role in quantifying gait complexity and automaticity rather than stability [52].

The primary objective of this systematic review is therefore to examine the use of long-term DE in nonlinear gait analysis, with a particular focus on its reinterpretation as a measure of gait complexity rather than stability. By synthesizing existing evidence, this review seeks to reconcile historical interpretations of long-term DE as a stability metric with emerging evidence highlighting its association with fractal dynamics and stride-to-stride fluctuation complexity. This synthesis may clarify the potential of long-term DE in the assessment of gait automaticity. A secondary objective is to address the substantial methodological heterogeneity observed across studies using long-term DE. This includes variations in measurement systems, signal processing techniques, and parameterization of Rosenstein's algorithm. By evaluating these methodological differences, we aim to identify best practices and propose standardized protocols for future research. Ultimately, this review aims not only to refine our understanding of long-term DE as a measure of gait complexity and automaticity, but also to facilitate more reliable and comparable applications in both research and clinical contexts.

Methods

Reporting and transparency

The systematic review followed the PRISMA guidelines to ensure transparency and comprehensive reporting [53]. However, it is important to note that the protocol for this review was not registered in PROSPERO or any other repository. Methodological reviews focusing on tool development or reporting practices without direct clinical applicability are typically excluded from PROSPERO registration requirements [54]. Despite the lack of formal registration, efforts were made to maintain transparency by publishing the protocol as a non-peer-reviewed preprint [55]. The readers are invited to consult the published protocol for further details about the review procedures.

Study eligibility criteria

The review included original peer-reviewed research articles focusing on human locomotion using the maximum Lyapunov exponent method with Rosenstein's algorithm to compute long-term DEs. Database searches applied no language restrictions to ensure comprehensive retrieval of potentially relevant literature. Eligible studies investigated gait in healthy participants or patients across all age groups and were published between 2000 and 2024. Observational, experimental, and intervention studies were included if they reported

sufficient methodological details on DE computation. Studies were excluded if they: focused on modes of locomotion other than walking (e.g., running or jumping), relied exclusively on short-term DEs or alternative analyses (e.g., Wolf's algorithm), were pure modeling or simulation studies lacking empirical human data, or did not adequately describe the DE computation parameters. Reviews, letters, editorials, and conference abstracts were also excluded. Full-text screening was restricted to articles published in English or French based on the linguistic competencies of the review team. Articles identified in languages other than English or French underwent title, abstract, and methods section screening using automated translation tools to assess eligibility against inclusion criteria.

To ensure the currency of the evidence base, an updated search was performed in January 2026 using the same databases and the identical search strategy, with the publication period restricted to 2024–2026. The updated screening applied the same two-phase selection process as the initial search. Two reviewers independently screened titles and abstracts, then assessed full texts for eligibility using the original inclusion criteria. Disagreements were resolved through discussion between the two reviewers.

Information sources and search strategy

We performed a systematic search using the Web of Science platform, including MEDLINE and the Web of Science Core Collection databases. The search strategy combines themes related to gait analysis, nonlinear dynamics, and Lyapunov exponents. Boolean operators and truncations were used to ensure comprehensive coverage. The search keywords are described in the supporting information (S1 Text).

Study selection process

The study selection process was conducted in two phases, the results of which are summarized in a PRISMA flowchart (Fig 1). Initially, two independent reviewers screened the titles and abstracts of all retrieved articles using broad inclusion criteria, as specific references to long-term divergence exponents were most often missing from abstracts. Articles potentially involving nonlinear gait analysis were advanced to the second phase, where their full texts were thoroughly reviewed to confirm the use of long-term divergence exponents calculated with Rosenstein's algorithm. Disagreements between reviewers were resolved through discussion, with arbitration by a third reviewer when necessary [55].

Data collection process and items

Main outcomes. Study outcomes were collected using standardized extraction tables as specified in the protocol [55], and the extracted results are available in the supporting information for detailed reference (S1 Text). For each included study, three main categories of information were extracted: study characteristics, methodological specifications, and outcomes related to short-term and long-term DEs. Study characteristics included publication year, authorship, sample size, participant demographics (age, health status), and research context (e.g., methodological studies or clinical evaluations). Methodological specifications included details on measurement methods (e.g., primary measurement system, signal type, sampling rate), walking conditions (overground or treadmill), signal processing procedures, and specific computational parameters for DEs (e.g., time delay, embedding dimension). Finally, the extracted findings focused on study objectives, conclusions, and interpretations of DE results.

To enable comparison of findings across studies, we computed relative percent change for each reported contrast, defined as the difference between the experimental and reference condition divided by the reference condition value, multiplied by 100 (Tables 1–3). This standardization was necessary because absolute DE values depend on the measurement system, signal type, and anatomical location, rendering raw values incommensurable across the diverse methodological approaches represented in the corpus. Expressing results as percent change removes this baseline dependency while preserving the direction and relative magnitude of effects within each study. Furthermore, percent change functions as a practical effect size metric, which is preferable to relying solely on statistical significance in this literature: many included studies used sample sizes that may be underpowered for detecting between-condition differences in long-term DE, and reporting effect magnitudes allows consistent identification of response patterns across studies regardless of whether individual comparisons reached significance. Finally, because most numerical values were extracted from published figures using a digital extraction tool (WebPlotDigitizer, https://automeris.io/), percent changes are more robust than absolute values to the small imprecision inherent in graphical data extraction.

Divergence curve characterization. In addition to extracting numerical outcomes, we systematically identified studies that published graphical representations of divergence curves. Divergence curves provide a visual representation of how the DEs are calculated. For each study that published divergence curves, we documented the figure location and extracted relevant information about the visual presentation of the method. For articles available through open-access licenses, we reproduced the published figures along with their complete legends. For studies behind paywalls where figure reproduction was not allowed,

we documented the figure legends and descriptions to catalog the types of divergence curve presentations.

Quality assessment

We assessed study quality using a structured evaluation framework across three domains. Each study was scored on a 15-point scale comprising Domain 1: analytical rigor, 8 points; Domain 2: outcome reporting, 4 points; and Domain 3: sample size, 3 points. Domain 1 assessed methodological transparency through four criteria: experimental protocol description, measurement system reporting, signal processing clarity, and algorithm application details. Domain 2 evaluated the quality of statistical reporting, specifically whether descriptive statistics were clearly reported, and whether appropriate statistical analyses with effect sizes were provided. Domain 3 considered sample size adequacy relative to each study's research question. This structured approach allowed for systematic comparison of methodological quality across the corpus while recognizing the exploratory nature of the field and the developmental stage of long-term DE methodology. Item-level and total quality scores for all included studies are provided in S3 Table (XLSX).

Synthesis method: meta-analysis

We identified five studies (contributing six independent study sets) that examined associations between long-term DEs and scaling exponents (DFA) [40,49,51,56,57] for synthesis via meta-analysis. Among the studies included in the systematic review, two publications [50,51] used overlapping samples: the first reported correlation results for 60 older adults, while the second extended the analysis to include 42 additional younger participants. Because the older adult sample in [50] is entirely subsumed within the larger dataset of [51], only the latter was included in the meta-analysis. These studies reported multiple correlation coefficients, corresponding to different measurements (accelerometer axes, body positions). Therefore, to synthesize their results, we performed a meta-analysis using a multilevel random-effects approach in accordance with current methodological recommendations for handling dependent effect sizes [58]. We implemented a correlated and hierarchical effects model using the metafor package in R [59]. This model accounts for both between-study heterogeneity and within-study correlation structures. We assumed within-study correlation of 0.8 between effect sizes from the same study, as the different correlation coefficients consisted of repeated measures in the same subjects. Within each study, repeated correlation coefficients were aggregated to a single study-level estimate

using a generalized least-squares average that respects the assumed within-study correlation. Robust variance estimation (RVE) with small-sample correction was applied using the clubSandwich package [60] to account for potential misspecification of the working model and to generate valid inferences despite the dependence structure of the data. Fisher's Z-transformed correlations were used for analysis and back-transformed to correlation coefficients for interpretation. We also estimated heterogeneity and summarized it with I^2. Results were visualized using forest plots displaying individual and overall estimates with their respective 95% confidence intervals.

Protocol deviations

We acknowledge five protocol deviations that occurred during the conduct of this review. First, the eligibility criteria for study inclusion originally specified a publication date range from 2001 to 2024. However, it became apparent that seminal articles relevant to the analysis of DE in nonlinear gait analysis were published as early as 2000. Consequently, the inclusion criteria were adjusted to include publications from 2000 to 2024. Second, although not explicitly stated in the original protocol, we decided during the data extraction phase to identify and report studies using the detrended fluctuation analysis (DFA) method in parallel to long-term DE. Third, we systematically documented divergence curve presentations to capture how these visualizations demonstrate the extraction of short-term and long-term DEs and reveal characteristics of the underlying gait dynamics. Fourth, the protocol specified a narrative quality assessment approach. However, upon initiating the assessment process, we recognized that a more structured framework would facilitate systematic evaluation and comparison across studies. We therefore developed a formal quality assessment grid with explicit scoring criteria across three methodological domains, enabling quantitative characterization of reporting quality while maintaining the focus on methodological transparency rather than risk of bias. Fifth, an updated search was conducted in January 2026 to capture studies published after the original search date of February 2024. This update, which was not anticipated in the protocol, was motivated by the extended time required to complete data extraction, quality assessment, manuscript preparation, and the publication process.

Results

Our systematic search retrieved 673 records from database queries and 3 additional records from references of other articles, yielding 676 unique records for title-abstract screening. Of these, 242 were excluded: 102 for non-relevant sources (proceedings, reviews), 51 for modeling studies, and 89 for irrelevant topics or non-walking gaits (running). Among the excluded records, 4 articles published in languages other than English or French (two Spanish, one German, one Chinese) underwent screening based on English abstract; none met the inclusion criteria. This left 434 full-text articles for eligibility assessment, all of which were in English. The full-text review resulted in the exclusion of 378 articles, mainly due to reporting only short-term DE. Other secondary reasons for exclusion were the use of alternative DE algorithms (Wolf) or other stability metrics.

An updated search conducted in January 2026 retrieved 94 additional records from the same databases. Following the same screening procedure, 36 records were excluded at the title-abstract stage (9 for non-relevant sources, 7 for modeling studies, and 20 for irrelevant topics or non-walking gaits), and 58 full-text articles were assessed for eligibility. Of these, 52 were excluded, and 6 new studies met the inclusion criteria, one of which contributed data to the meta-analysis.

In total, the combined searches yielded 770 records, of which 62 studies were included in the outcome synthesis and 5 in the meta-analysis (Fig 1).

Methodological outcomes

Study characteristics, methodological scope, and sample sizes. Fig 2 summarizes the characteristics of the 62 included studies, based on the data provided in the supporting information (Table B in S1 Text). Panel a illustrates the annual number of publications from 2000 to 2026, showing a consistent upward trend starting in 2008, followed by sustained activity through the mid-2010s and a resurgence of four publications in 2024. Panel b categorizes the type of research by participant category. Exact definitions for each research type are provided in supporting information (Table A in S1 Text). In terms of populations studied, studies involving only healthy adults predominate, with 42 studies. Within this category, most studies compared gait patterns under different experimental conditions and aimed to develop, validate, or assess the reliability of DE methods. Research on age-related gait changes and fall risk comprises 14 studies, while the remaining 6 studies investigated gait patterns associated with other pathologies. Only a few studies have used long-term DE for clinical investigations (4). Only one study was interventional and longitudinal, i.e. it analyzed the gait patterns of patients before and after surgery. Aside from studies aimed at

predicting fall risk, we identified no studies that evaluated the use of long-term DE as a diagnostic tool for detecting gait abnormalities. Panel c shows that across the study samples, the median sample size was 23 participants (interquartile range 13-36), with most studies enrolling fewer than 50 individuals; only a handful exceeded 100, and the full range extended from 9 to 187 participants, indicating a generally modest sample size framework in the field. Note that one article was counted as two different studies because it reported two separate experiments with two different groups of participants.

Experimental settings and sensor modalities. Per-study methodological outcomes for all 62 studies are provided in Table C in S1 Text. Measurement environments were overwhelmingly laboratory-bound (Fig 3a). Forty-two of the 62 studies (68%) acquired data on an instrumented treadmill in an indoor laboratory setting, typically in conjunction with 3-D motion-capture systems. Thirteen studies (21%) relied solely on overground walking, but even here the majority preserved a controlled indoor context (indoor circuits or short walkways). Only three overground protocols departed from this paradigm by recording prolonged outdoor or quasi-natural bouts. Six further articles (10%) combined treadmill and overground sessions within the same experiment, whereas one study (2%) captured free-living gait during a week-long monitoring campaign. Instrumentation choices mirrored these environmental constraints (Fig 3b). Optical motion-capture was used in 31 of the 62 studies (50%). Wearable inertial measurement units (IMUs) ranked second, appearing in 19 studies (31%). Dedicated instrumented treadmills—equipped with embedded force or pressure arrays—were the unique measurement device in 3 studies (5%). Although single-sensor protocols remained the norm, a minority of nine studies (15%) employed hybrid instrumentation—most commonly electrogoniometers coupled with a trunk accelerometer (4/9, 44%), followed by optical motion-capture combined with body-worn IMUs (3/9, 33%) and motion capture and surface electromyography (EMG, 2/9, 22%). Most studies prioritized measurement at the trunk level (Fig 3c): 34 of the 62 studies (55%) analyzed pelvis, lumbar or thoracic spine kinematics. Lower-limb–only set-ups (hip, knee, ankle kinematics or EMG) appeared in 11 studies (18%). Eleven studies (18%) combined trunk and lower-limb measurements within the same experiment. Foot-pressure–only configurations were rare (3/62, 5%), corresponding to the instrumented-treadmill studies. Other designs include head, upper-limb, or composite whole-body recordings (3/62, 5%).

Measurement lengths. Because Rosenstein's algorithm requires a single continuous signal, we extracted the duration corresponding to one analyzed walking bout per experimental condition, rather than aggregating multiple bouts. Reporting practices across the included studies varied substantially: approximately one-third specified bout length

exclusively in minutes, another third exclusively in strides, and the remainder provided both units. Due to this methodological heterogeneity, direct quantitative synthesis of durations was not feasible. We therefore adopted an ordinal categorization to harmonize duration data into four empirically defined classes (Fig 4). The threshold for the "very short" category (≤ 2 min or ≤ 100 strides) was anchored to the median cadence observed in studies reporting both units (~55 strides·min^{-1}), yielding approximately 100 strides over two minutes—a practical lower limit below which long-term DE estimates may become unreliable [61]. The upper limit for the "short" category (> 2–5 min or 101–250 strides) aligns with the most commonly adopted experimental protocol across studies (3–5 minutes of treadmill or overground walking). The "moderate" category (> 5–10 min or 251–500 strides) encompasses studies employing somewhat longer walking bouts, while the "long" category (> 10 min or > 500 strides) includes the few studies focused on endurance-oriented walking conditions. Applying this ordinal scale to the extracted data resulted in the following distribution: "short" (30 studies; 48%), "very short" (22 studies; 35%), "moderate" (6 studies; 10%), and "long" (4 studies; 6%). Finally, five studies computed DE from averaged signals of multiple short walking bouts within the same experimental condition; for these, we categorized duration based on a single constituent bout, although averaging multiple bouts may enhance the reliability of DE estimates [61].

Signal processing. Signal sources driving the computation of long-term DEs were clustered into distinct categories (Fig 5a). Half of the studies (31 out of 62, 50%) utilized kinematic data exclusively, acquired via optical motion capture or electrogoniometers. Within these kinematic studies, angular joint trajectories (joint angles) were predominant, representing most cases. Additionally, linear positional data (e.g., cervical or thoracic spine positions) and derived velocity data were also commonly used. Accelerometric data from inertial sensors constituted the second most frequent signal type, with 18 studies (29%) employing linear accelerations exclusively. Only two studies (3%) incorporated combined inertial measurements of linear acceleration and angular velocity from gyroscopes. No studies used gyroscopic (angular velocity) signals alone. Three studies (5%) derived DEs from COP trajectories obtained from instrumented treadmills. Surface EMG voltage signals alone were used in a single study (2%). Mixed signal sources appeared infrequently: six studies (10%) combined joint-angle kinematics with trunk acceleration, and one study (2%) combined whole-body kinematics with multi-channel EMG. Reporting of raw-signal filtering was inconsistent across the corpus. Twenty-one of the 62 studies (34%) explicitly described applying a digital filter before computing the long-term DE, almost always a zero-lag low-pass Butterworth between 8 Hz and 20 Hz. Twenty-two papers (35%) stated unambiguously

that no filtering was performed. For the remaining nineteen studies (31%) no filtering procedure was mentioned, making it probable—though not certain—that the raw signals were analyzed without additional pre-processing.

The next steps in DE computation typically involve time normalization of raw data. The rationale is to maintain a consistent number of steps and data points across participants and conditions. This approach prevents potential bias arising from variations in DE values associated with sample length [61]. Panel b of Fig 5 shows the handling of temporal resampling. Explicit time normalization—typically realized as linear or spline interpolation to a constant number of samples—was reported in 25 of the 62 investigations (40%). Three studies (5%) used an alternative method consisting in resampling separately each stride to a fixed number of samples. In contrast, 32 papers (52%) gave no information on resampling, very likely using raw, unequal-length signals, and / or signals with a variable number of steps. The remaining two articles (3%) explicitly did not perform time normalization.

Algorithms. The subsequent processing involves reconstructing the state-space on which DE is computed. Among the 62 studies, 54 (87%) used the classical Takens approach (see Appendix A in [40]), generating a d_E-dimensional trajectory from a single signal and its T-delayed copies [62]. By contrast, 8 studies (13%) constructed ad-hoc, multi-channel attractors by concatenating measured variables—typically linear plus angular trunk kinematics, or acceleration combined with gyroscope signals—before any delay embedding. Within the subset of studies using the classical approach (54 studies), the tuning procedure for the delay T and embedding dimension d_E was reported heterogeneously: 34 of 54 studies (63%) applied the standard pair of average mutual information (AMI) for T and global false-nearest-neighbors (GFNN) for d_E, seven (13%) substituted the autocorrelation criterion for T while still using GFNN, and the remaining thirteen (24%) gave no explicit or unclear justification for either parameter.

Finally, after reconstructing the state-space and computing divergence curves, the long-term DE is determined by performing a linear fit on a specific segment of the divergence curve [40]. The interval used for this linear fit was initially defined in seminal studies by Dingwell et al. [41,63]. Most studies (55 out of 62, 89%) adopted this canonical interval, 4 to 10 strides. Among the seven studies (11%) that selected alternative ranges, four used intervals between 0.5 and 10 strides (including one study that explored multiple sub-intervals), while three recent studies employed ranges shifted toward later strides (5–12 and 6–8 strides), two of which evaluated multiple range/length combinations to optimize responsiveness and reliability of the long-term DE estimate.

Divergence curve visualization. Divergence curves—graphical representations plotting the average logarithmic divergence of neighboring trajectories against time—demonstrate how short-term and long-term DEs are extracted as linear slopes from distinct temporal regions. In addition, their shape may allow visualizing when divergence ceases to grow as trajectories reach bounded regions of phase space. Among the 62 included studies, 30 (48%) published divergence curves in their figures. We systematically categorized the purpose for which divergence curves were presented across these studies. Eighteen studies (60%) used divergence curves exclusively for methodological illustration, displaying representative curves to demonstrate the calculation procedure without comparing experimental conditions. Seven studies (23%) employed divergence curves exclusively as comparative results, presenting curves to highlight differences between groups or experimental conditions. Five studies (17%) included both types of presentations, incorporating separate figures for methodological illustration and comparative outcomes. Among studies presenting comparative results, typical comparisons included different walking speeds, perturbation versus control conditions, young versus older adults, and with versus without external cueing. The compiled divergence curves and associated figure descriptions are provided in S2 Figures.

Quality Assessment

Quality assessment was completed using the three-domain scoring framework (maximum 15 points). Overall methodological quality was moderate to high, with a mean score of 11.56 (SD = 1.78, range 8-15) out of 15 possible points. Twenty-seven studies (44%) achieved high quality (12-15 points), while 35 studies (56%) demonstrated moderate quality (8-11 points). No studies scored in the low-quality range (0-7 points). Complete quality assessment scores for all studies are provided in S3 Table.

Domain-specific analysis revealed distinct patterns across methodological areas. Domain 1 (analytical rigor) demonstrated the strongest performance, with studies averaging 7.60 ± 0.66 out of 8 points (95%), and 43 studies (69%) achieving perfect scores (8/8). This indicates robust methodological rigor in experimental protocols, measurement systems, signal processing, and algorithm implementation across the included literature. In contrast, Domain 2 (outcome reporting) showed moderate performance (mean 2.31 ± 1.15 out of 4 points; 58%), with only 13 studies (21%) achieving full credit for comprehensive statistical reporting. Domain 3 (sample size) similarly revealed limitations, averaging 1.66 ± 0.81 out of 3 points, with most studies (40%) demonstrating moderate sample size justification and only 10 studies (16%) providing adequate sample size justification.

Long-term DE results

Perturbation studies. Among the 62 studies, 14 studies experimentally destabilized gait using a range of perturbation modalities, including mechanical disturbances (e.g., compliant surfaces, outdoor surfaces, segmental loads, unexpected slips, platform instability), sensory manipulations (e.g., galvanic vestibular stimulation, blindfolding, optical-flow distortions, foot desensitization), or cognitive disturbance (dual-tasking). These experimental paradigms are particularly relevant for testing the hypothesis that long-term DE reflects the degree of gait automaticity. When gait is perturbed or destabilized, it is reasonable to assume that individuals allocate greater attentional resources to gait control in order to manage the imposed challenges. However, none of the 14 perturbation studies explicitly tested this hypothesis, as they interpreted long-term DE in the same manner as short-term DE—that is, as a direct index of gait stability sensitive to the magnitude of external perturbation. We summarize the results of the retained perturbation studies in Table 1. Because long-term DE is sensitive to signal length (number of data points or strides), valid comparisons require similar data lengths across conditions. In treadmill perturbation paradigms—where walking speed is held constant between baseline and challenge—data length remains comparable even without time normalization. Therefore, we excluded studies with very-short bouts or overground designs in which walking speed (and thus signal length) differed between conditions, retaining only treadmill experiments with sufficient bout duration and matched speeds. We also excluded one study that used stride-level temporal normalization [64] which suppresses long-term divergence [52] and renders values non-comparable to those obtained with time-based resampling. Note that the study by Manor et al. analyzed a reduced number of consecutive strides (45), but reported results averaged across three different walking speeds, which may enhance the reliability of the DE assessment.

Table 1. Results of the perturbation studies

Study	Measurement method	Perturbation	Short-term DE change with perturbation	Long-term DE change with perturbation
Manor et al. ID 15 [65]	2D motion capture, sagittal plane hip, knee, and ankle angles.	Ice-induced plantar desensitization	**+8%**	**+44%**

McAndrew et al. ID 22 [66]	3D motion capture, 3D velocity of the cervical spine.	Mechanical and visual	Mechanical: **+10% to +72%** Visual: **+3% to +86%**	Visual: **-7% to -51%** Mechanical: **+5% to -38%**
Van Schooten et al. ID 26 [48]	3D inertial sensor, accelerations and angular velocities of the lower back.	Vestibular	**+8% to +10%**	**-4% to -20%**
Sinitksi et al. ID 30 [67]	3D motion capture, mediolateral velocity of the cervical spine.	Mechanical and visual	Mechanical: **+30% to +34%** Visual: **+19% to +34%**	Mechanical: **-27% to -30%** Visual: **-9% to -35%**
Rábago et al. ID 42 [68]	3D motion capture, mediolateral velocity of the cervical spine.	Cognitive, mechanical, and visual	Cognitive: **-3% to 0%** Mechanical: **+8% to +28%** Visual: **+10% to +30%**	Cognitive: **+6% to -3%** Mechanical: **-16% to -39%** Visual: **-17% to -41%**
Reynard and Terrier ID 43 [69]	3D accelerometer. Mediolateral acceleration of the upper trunk (sternum).	Vision removal	**-2% to +7%**	**0% to -20%**
Franz et al. ID 41 [33]	3D motion capture. Right and left heel positions relative to the sacrum position.	Visual	Young: **+6%** Older: **+22%**	Young: **-13%** Older: **-17%**
Qiao et al. ID 47 [70]	3D motion capture, 3D velocity of the cervical spine.	Visual	Healthy young controls: **+19% to +24%** Elderly non-fallers: **+20% to +23%** Elderly fallers: **+19% to + 28%**	Healthy young controls: **-7% to -32%** Elderly non-fallers: **+2% to -49%** Elderly fallers: **-7% to -44%**

Perturbation types included sensory manipulation (ice-induced plantar desensitization), mechanical disturbances (platform translations), visual alterations (moving room, optic flow), vestibular stimulation (galvanic stimulation), and cognitive tasks (dual-task paradigms). Percentage changes represent relative differences between perturbation and baseline conditions. Positive values indicate increased divergence exponent (DE) while negative values indicate decreased DE. Ranges reflect different measurement axes, experimental groups, or perturbation levels within individual studies. ID refers to study identification number in the systematic review database (S1 text in supplementary material).

Study designs differed in the nature and number of perturbation conditions, which explains why we report ranges rather than single effect sizes. For most studies, numeric values were extracted from figures using WebPlotDigitizer (https://automeris.io/). Manor et al. [65] desensitized the plantar arch using ice to diminish sensory feedback from the feet,

anticipating subsequent gait destabilization. Unlike other perturbation studies that analyzed trunk motion, Manor et al. focused on lower-limb kinematics. Their findings indicated an increase in both short-term (+8%) and long-term (+44%) DE values, with results aggregated across different walking speeds and joints. In McAndrew et al. [66], participants experienced both mediolateral and anteroposterior platform oscillations and visual-field perturbations while DE was computed along the AP, ML, and vertical axes, yielding short-term DE increases of 10–72% (mechanical) and 3–86% (visual) and long-term DE decreases of 5–38% and 7–51%, respectively. Sinitski et al. [67] systematically varied the amplitude of both mechanical and visual perturbations, observing short-term DE rises of 30–34% (mechanical) and 19–34% (visual) alongside long-term DE drops of 27–30% and 9–35%, respectively. Rábago et al. [68] compared cognitive, mechanical, and visual perturbations: cognitive load had negligible effects on short-term DE (−3% to 0%) with near-zero long-term changes (−3% to +6%), while both mechanical and visual perturbations produced a similar pattern of increased short-term DE (+8% to +30%) accompanied by decreased long-term DE (−16% to −41%). Van Schooten et al. [48] and Reynard & Terrier [69] compared walking at multiple speeds under galvanic vestibular stimulation and blindfolded conditions. They found short-term DE increases of 8–10% (GVS) and –2 to +7% (blindfold), and long-term DE decreases of 4–20% and 0–20%, respectively. Franz et al. [33] analyzed younger and older subjects, reporting short-term DE increases of 6% (young) and 22% (older), and long-term DE decreases of 13% (young) and 17% (older) under visual perturbation; they also found a reduction in the scaling exponent reflecting gait complexity (DFA method) by 6% in young and 26% in older participants. The study by Qiao et al. assessed the effects of varying intensities of optical-flow distortions on young and older adults. Here again, opposite responses were observed between short-term DE, which increased (+19% to +28%), and long-term DE, which mostly decreased (+2% to -48%). Across all studies reviewed, the majority of perturbation-induced changes were statistically significant ($p < 0.05$), though certain exceptions occurred, notably for long-term DE measures that exhibited greater inter-individual variability.

Between-subject studies. Fifteen between-subject studies compared cohorts with presumed differences in gait stability and potentially in gait automaticity, including clinical groups (e.g. diabetic or peripheral neuropathy, knee osteoarthritis, chronic pain, fallers vs non-fallers) and demographic cohorts (healthy older vs younger adults, inexperienced vs experienced military cadets). However, most of these studies based their DE analysis on too short walking bouts or on non-time normalized signals. After careful evaluation of study methodologies, eight studies met the inclusion criteria for a valid comparison (Table 2).

585 Although Terrier et al. analyzed very short walking bouts (35 strides, 1 min), they
586 aggregated numerous bouts collected over one week under free-living conditions (on
587 average 77 bouts in patients and 34 in healthy controls). This aggregation enhances the
588 reliability of the average DE values, mitigating concerns related to short individual signal
589 durations.
590

591 **Table 2. Results of the between-subject studies**

Study	Measurement method	Population	Short-term DE difference with controls	Long-term DE difference with controls
Manor et al. ID 9 [71]	2D motion capture. Sagittal plane joint angles for hip, knee and ankle	Patients with all-cause peripheral neuropathy	**+4%**	**-4%**
Franz et al. ID 41 [33]	3D motion capture. Right and left heel positions relative to the sacrum position.	Healthy older adults	Normal walking: **+1%** Perturbation: **+15%**	Normal walking: **-10%** Perturbation: **-15%**
Terrier et al. ID 45 [72]	3D acceleration, right hip.	Patients with chronic musculoskeletal pain of the lower limbs.	**+4%**	**-19%**
Bizovska et al. ID 46 [73]	3D acceleration of the lumbar spine	Older healthy female subjects	Overground: **+4% to +18%** Treadmill: **+3% to +25%**	Overground: **+56% to -11%** Treadmill: **+22% to -8%**
Qiao et al. ID 47 [70]	3D motion capture, 3D velocity of the cervical spine.	Older non-faller adults and older faller adults	Young controls vs. older non-fallers: **+6% to +13%** Young controls vs older fallers: **+14% to +19%**	Young controls vs. older non-fallers: **+37% to -45%** Young controls vs older fallers: **+49% to -36%**
Bizovska et al. ID 49 [74]	3D acceleration of the trunk	Elderly fallers and frequent fallers	Older controls vs. fallers: **+7% to +8%** Older controls vs frequent fallers: **-14% to +11%**	Older controls vs. fallers: **+8% to -3%** Older controls vs frequent fallers: **-7% to -18%**
Matinazad et al. ID 50 [75]	3D motion capture, sagittal angle of the knee	Older female patients with knee osteoarthritis	**+41% to +43%**	**-43% to -85%**
Piergiovanni and Terrier. ID S4 [51]	3D acceleration of the trunk	Healthy older adults	Normal walking: **+8%** Metronome walking: **+12%**	Normal walking: **-15% to -21%** Metronome walking: **-29% to -33%**

592 *Percentage differences in short-term and long-term divergence exponents (DE) between*
593 *clinical/aging populations and control groups. Populations included patients with all-cause*
594 *peripheral neuropathy, chronic musculoskeletal pain of lower limbs, knee osteoarthritis, and*

older adults (non-fallers, single fallers, and frequent fallers). Positive values indicate higher DE in the comparison group relative to controls, while negative values indicate lower DE. Ranges reflect different measurement axes, walking conditions, or perturbation levels within individual studies. ID refers to study identification number in the systematic review database (S1 text in supplementary material).

Study methodologies differed in signal processing approaches and reported outcomes. Manor et al. [71] averaged their analyses across joint angles and walking speeds, reporting a single DE value per group (patients with peripheral neuropathy vs. healthy controls). Similarly, Terrier et al. [72] computed the vector magnitude (norm) from three axes of hip acceleration, yielding a single representative DE per group (chronic musculoskeletal pain patients vs. controls). Using 3D motion capture during treadmill walking, Franz et al. [33] examined trunk motion responses to visual perturbations. Similarly, the visual perturbation study of Qiao et al. [70] examined cervical spine velocity under three increasing perturbation levels. Bizovska et al. [73] reported separate DE outcomes for each axis of a 3D accelerometer under two distinct walking conditions (overground vs. treadmill), resulting in a range of observed differences between older and younger healthy females. Similarly, Piergiovanni & Terrier [51] reported long-term DE across three axes of a lower-back accelerometer under normal and metronome walking conditions, comparing older to younger adults. In their second study, Bizovska et al. also analyzed DEs across accelerometer axes, comparing a control group of healthy older adults with groups of older individuals who fell infrequently or frequently. Finally, Matinazad et al. studied knee kinematics at two different speeds.

Short-term DE demonstrated a pattern of limited increases in most populations compared to their respective controls, though with heterogeneity in magnitude. The most pronounced differences were observed in patients with knee osteoarthritis (41-43% higher). Clinical populations with neuropathy and chronic musculoskeletal pain showed consistent increases compared to controls (+4%). Regarding age comparison studies, Franz et al. reported minimal differences during normal walking (1% higher in older adults as compared to younger controls) but substantially larger increases under visual perturbations (15% higher in older adults). Piergiovanni & Terrier [51] found a similar pattern, with 12% higher short-term DE in older adults during normal walking and 8% higher during metronome walking. Qiao et al. [70] found higher short-term DE in older adults versus younger adults, with this difference amplified by perturbation intensity, especially in older fallers. Bizovska et al. [73] found short-term DE differences ranging from 4% to 18% higher in older compared to younger females during overground walking, and from 3% to 25% higher during treadmill

630 walking. However, it is worth noting that the older group had a mean age of only 57.5 years
631 (±4.8), representing middle-aged rather than elderly adults. In their second study focused
632 on fall risk, Bizovska et al. [74] observed short-term DE increases of 7-8% when comparing
633 elderly non-fallers to single fallers, while comparisons between non-fallers and frequent
634 fallers showed highly variable results ranging from 14% decreases to 11% increases
635 depending on measurement axis.
636 In contrast to the relatively consistent short-term increases, long-term DE exhibited
637 heterogeneous patterns. The most substantial reductions were observed in clinical
638 populations: patients with knee osteoarthritis showed decreases ranging from 43% to 85%
639 compared to controls, while those with chronic musculoskeletal pain demonstrated 19%
640 reductions. Neuropathy patients showed minimal reductions of 4%. Franz et al. [33]
641 reported consistent long-term DE reductions in older adults during both normal walking (-
642 10%) and under visual perturbations (-15%). Piergiovanni & Terrier [51] found comparable
643 reductions during normal walking (−15% to −21% across axes), which were further
644 amplified during metronome walking (−29% to −33%). Despite reporting heterogeneous
645 results, Qiao et al.'s figures [70] show a trend toward lower long-term DE (−36%) in older
646 participants under high visual perturbation. The two studies by Bizovska et al. revealed
647 higher variability depending on measurement axis and experimental conditions. In their age-
648 comparison study [73], long-term DE differences ranged from increases of 56% to decreases
649 of 11% when comparing older and younger females during overground walking, and from
650 increases of 22% to decreases of 8% during treadmill walking. Their fall-risk study [74]
651 showed long-term DE differences ranging from increases of 8% to decreases of 3% when
652 comparing non-fallers to single fallers, and from decreases of 7% to 18% when comparing
653 non-fallers to frequent fallers. Overall, three of the four age-comparison studies showed
654 consistent long-term DE reductions in older adults; only the Bizovska data, based on a
655 relatively young older cohort (mean age 57.5 years), diverged from this pattern.

656

657 **Table 3. Results of the externally paced walking studies**

658

Study	Measurement method	Cueing	Short-term DE change with cueing	Long-term DE change with cueing
Dingwell & Cusumano ID 3 [63]	Electrogoniometers (hip, knee, ankle) + sternum tri-axial accelerometer.	Treadmill speed constraint (vs overground).	Trunk acceleration: **-11% to -13%** Lower limb kinematics:	Trunk acceleration: **-11% to +12%** Lower limb kinematics: **-11% to -20%**

			-11% to -16%	
Terrier & Dériaz ID 25 [56]	3D accelerometer on lower back (lumbar), trunk accelerations.	Treadmill speed constraint (vs overground).	**-8% to -9%**	**-5% to -18%**
Sejdić et al. ID 29 [47]	Force-sensitive insole (heel strikes) + lower-back 3D accelerometer, trunk accelerations.	Rhythmic sensory cues during overground walking, step duration constraint: auditory, visual, and tactile.	Visual / step duration: **-4% to +3%** Auditory / step duration: **-11% to -4%** Tactile / step duration: **-8% to 0%** Combined / step duration: **-12% to -8%**	Visual / step duration**:** **-10% to -18%** Auditory / step duration: **-26% to -40%** Tactile / step duration: **-15% to -23%** Combined / step duration: **-34% to -43%**
Terrier & Dériaz ID 35 [40]	Instrumented treadmill (foot-pressure array); center-of-pressure trajectory.	Rhythmic auditory cueing (metronome), step duration constraint, during treadmill walking.	**-6% to +7%**	**-61% to -86%**
Bizovska et al. ID 46 [73]	Lumbar 3D accelerometer, trunk accelerations.	Treadmill speed constraint vs overground.	Young: **-19% to -24%** Older: **-16% to -21%**	Young: **+24% to +156%** Older: **0% to + 100%**
Terrier ID 51 [57]	Instrumented treadmill with force platform; center-of-pressure trajectory.	Visual (step length constraint) and auditory (step duration constraint) synchronization cues, during treadmill walking.	Visual / step length: **-4% to -2%** Auditory / step duration: **-3% to + 14%**	Visual / step length: **-37% to -66%** Auditory / step duration: **-43% to -78%**
Piergiovanni and Terrier. ID S4 [51]	3D acceleration of the trunk	Rhythmic auditory cueing (metronome), step duration constraint, during overground walking.	Young: **+1%** Older: **-3%**	Young: **−13% to −15%** Older: **−26% to −30%**

Percentage changes in short-term and long-term divergence exponents (DE) during externally paced walking compared to self-paced conditions. Interventions included treadmill speed constraint (imposing constant velocity versus self-paced overground walking) and rhythmic sensory cueing (auditory metronome, visual pacing, tactile stimulation). Negative values indicate reduced DE with external pacing, while positive values indicate increased DE. Ranges reflect different cueing modalities, measurement signals, or experimental conditions within studies. Combined cueing indicates simultaneous presentation of multiple sensory modalities. ID refers to study identification number in the systematic review database (S1 Text in supplementary materials).

External cueing studies. Among the 62 studies, seven within-subject experiments examined either externally imposed gait synchronization or a fixed-speed treadmill

constraint. As these interventions likely shift the balance between automatic and voluntary control of walking [76], they enable evaluation of the sensitivity of long-term DE to variations in gait automaticity. Three studies implemented synchronization cues: Sejdić et al. [47] applied visual, auditory, tactile, and combined step-duration cues during overground walking using heel-strike detection with a force-sensitive insole and lumbar accelerometry; Terrier & Dériaz [40] and Terrier [57] imposed metronome-based step-duration or visual step-length targets during treadmill walking while recording center-of-pressure trajectories on an instrumented treadmill. Piergiovanni & Terrier [51] applied auditory metronome cueing during overground walking in both younger and older adults, recording trunk accelerations from a lower-back accelerometer. Three studies compared constant-speed treadmill with unconstrained overground walking: Dingwell & Cusumano [63] recorded trunk accelerations (sternum) and lower-limb joint angles via electrogoniometers, and Terrier & Dériaz [56] used lower-back accelerometry; Bizovska et al. [73] extended this comparison in young and older women with lumbar accelerometry. Trial durations typically ranged from 3 to 10 min per condition (yielding ~100–500 strides), with speed held constant in cueing paradigms; overground–treadmill contrasts occasionally differed in bout duration or speed (e.g., 5 min overground vs 3 min treadmill in [73]).

Across studies that imposed external synchronization cues (auditory, visual, tactile), short-term divergence exponents (DEs) changed little or inconsistently (approximately −11% to +14%), whereas long-term DEs showed large, systematic reductions (≈ −10% to −86%). In the overground study using rhythmic sensory cues [47], short-term DEs were minimally affected (visual: −4% to +3%; tactile: −8% to 0%) or modestly decreased under auditory and combined cueing (auditory: −11% to −4%; combined: −12% to −8%). By contrast, long-term DEs fell substantially (visual: −10% to −18%; tactile: −15% to −23%; auditory: −26% to −40%; combined: −34% to −43%). Similarly, a substantial decrease in gait complexity measured from force-sensitive insoles (scaling exponent, DFA) was observed (-50% to -13%). In the second overground cueing study, Piergiovanni & Terrier [51] observed the same dissociation: auditory metronome cueing left short-term DE essentially unchanged (young: +1%; older: −3%) while producing substantial long-term DE reductions that were more pronounced in older adults (−26% to −30%) than in younger participants (−13% to −15%). Two treadmill experiments that enforced step timing/length converged on the same pattern. Using an instrumented treadmill and center-of-pressure (COP) trajectories, Terrier & Dériaz [40] reported very large long-term DE decreases (−61% to −86%) with only small, direction-dependent short-term changes (−6% to +7%). In a mixed visual–auditory cueing paradigm [57], short-term DEs remained near baseline (visual/step-length: −4% to −2%;

auditory/step-duration: −3% to +14%), whereas long-term DEs dropped markedly for both constraints (visual/step-length: −37% to −66%; auditory/step-duration: −43% to −78%).

By contrast, comparing treadmill speed constraint with unconstrained overground walking yielded consistent short-term DE reductions but heterogeneous long-term responses across datasets and signal types. In three studies that contrasted fixed-speed treadmill with overground walking, short-term DEs were lower on the treadmill in trunk accelerations ([63]: −11% to −13%, [56]: −8% to −9%, and [73] −16% to −24%), and lower-limb kinematics ([63]: −11% to −16%). Long-term DEs, however, varied: they mostly decreased in two datasets ([63]: trunk −11% to +12%, lower-limb −11% to −20%; [56]: −5% to −18%), but substantially increased in [73] (+24% to +156%).

Meta-analysis

We synthesized six independent datasets reporting the association between long-term DE and the scaling exponent from DFA. Measurement modalities spanned instrumented treadmills with center-of-pressure trajectories, trunk accelerometry, and segmental kinematics of the head and lower limbs; several experiments also varied speed (slow, preferred, fast) or imposed synchronization tasks. One publication contributed two distinct datasets because the same participants walked in separate overground and treadmill experiments [56]; these were treated as independent due to the large contextual difference between conditions. Where articles reported axis- or segment-specific associations, we retained those dependent effects (e.g., anteroposterior, mediolateral, vertical, or head vs ankle). This yielded 14 dependent correlations across the six datasets (k=14), while speed manipulations within a given protocol were not counted as separate effects when correlations were reported on the pooled observations [49].

Using a correlated-effects random-effects model with cluster-robust variance estimation, Fisher's Z-transformed correlations were pooled and then back-transformed for interpretation. The overall association was positive, with a pooled correlation of 0.64 (95% CI 0.34 to 0.82) based on 14 effects (N=582 condition-level observations from 209 participants). Heterogeneity was considerable ($I^2 \approx 82\%$), with estimated between-study variance $\tau^2 \approx 0.22$ and within-study variance $\omega^2 \approx 0.011$. Leave-one-out influence diagnostics showed that the pooled correlation remained stable in magnitude and direction. Re-estimating the model after omitting each dataset in turn yielded pooled r values between 0.57 and 0.71. Across all leave-one-out fits, the 95% confidence intervals stayed positive,

739 and the standard error varied only modestly (0.17–0.24). The forest plot summarizes study-
740 level estimates and the model summary (Fig 6).
741

742 # Discussion

743 This systematic review synthesized evidence from 62 studies to critically examine the use of
744 long-term divergence exponents (DE) computed by Rosenstein's algorithm in human gait
745 analysis, with particular focus on its reinterpretation as a measure of gait complexity and
746 automaticity rather than stability.

747 ## Summary of main findings

748 Our meta-analysis (Fig 6) provides empirical validation for the theoretical reconceptualization
749 of long-term DE as a complexity measure. The moderate-to-substantial positive correlation (r
750 = 0.64) between long-term DE and the DFA scaling exponent demonstrates that these
751 metrics capture a shared underlying construct—the fractal-like correlation structure of stride-
752 to-stride fluctuations. This association persisted robustly across heterogeneous measurement
753 contexts. Two additional studies [33,47] showed parallel reductions in both long-term DE
754 and scaling exponent during visual perturbation and external cueing, qualitatively supporting
755 the meta-analytic correlation despite not reporting correlation statistics.
756 The differential responsiveness of short-term and long-term DE across experimental designs
757 provides further support for their distinct physiological interpretations. Perturbation studies
758 (Table 1) revealed a consistent dissociation: destabilizing interventions systematically
759 increased short-term DE while decreasing long-term DE, confirming that these indices may
760 reflect separable control processes underlying gait. This pattern directly contradicts the
761 interpretation of both measures as stability metrics, since a truly destabilizing intervention
762 should increase both if they assessed the same construct. Between-subject comparisons
763 (Table 2) yielded more heterogeneous patterns, with clinical populations sometimes
764 exhibiting paradoxical long-term DE responses (both increases and decreases relative to
765 controls), further highlighting its dissociation from stability measures that should consistently
766 elevate with pathology. Finally, external cueing studies (Table 3) demonstrated that
767 interventions requiring conscious temporal or spatial coordination—which engage voluntary
768 control without perturbing balance—produced strong long-term DE reductions (up to −86%)
769 with minimal short-term DE changes, as anticipated when gait automaticity is reduced
770 without compromising stability. This selective sensitivity to sensory cueing, combined with

the parallel reductions observed in DFA scaling exponents [76], further support long-term DE as a marker of gait complexity rather than stability.

Long-term divergence saturation as a signature of reduced gait complexity

Divergence in strange attractors. Rosenstein's algorithm was originally developed to estimate the maximal Lyapunov exponent in deterministic chaotic systems. The method's computational validity was established using well-characterized strange attractors, including the three-dimensional Lorenz system—a nonlinear dynamical system defined by three coupled differential equations that exhibits sensitive dependence on initial conditions [77]. When applied to such chaotic systems, the logarithmic divergence curve exhibits characteristic behavior: an initial linear region from which the maximal Lyapunov exponent is computed, followed by a plateau where further divergence ceases (see Appendix A in [40]). This saturation occurs because strange attractors, despite their complex internal dynamics, occupy a bounded region of phase space, constraining the maximum separation between initially neighboring trajectories.

Divergence in gait dynamics attractors. Dynamical systems models suggest that human gait may exhibit periodic limit cycle attractor dynamics with stochastic perturbations, rather than strange attractors characteristic of deterministic chaos [15,38,78], a characterization supported by both discrete stride interval analyses [79] and continuous signal phase space reconstruction [78]. This distinction justifies the preference for "divergence exponent" over "Lyapunov exponent" when describing gait dynamics [15]. The divergence curves obtained from gait attractors reflect departure from purely chaotic systems: they typically display an initial exponential divergence phase followed by gradual saturation over subsequent strides but often lack the abrupt plateau characteristic of purely deterministic chaotic systems with bounded strange attractors (S2 Figures).

Fractal stride patterns and phase space exploration. A gait attractor constructed using Takens theorem or an alternative approach represents the repertoire of available coordination patterns—the different kinematic combinations and their temporal organization—with phase-space exploration describing how trajectories navigate through this repertoire across consecutive gait cycles [80]. The temporal correlation structure of consecutive strides likely modulates how trajectories explore the phase space, thereby influencing divergence curve saturation and hence long-term DE values. Low-frequency variations distributed across dozens of consecutive strides may create drift through multiple regions of the attractor, delaying the saturation process and yielding higher long-term DE

values. This mechanism may explain the observed correlation between long-term DE and DFA scaling exponents. In support of this hypothesis, Terrier and Reynard [52] manipulated stride interval correlation structure while preserving acceleration signal shape, demonstrating that long-term DE scaled systematically with temporal complexity while short-term DE remained unchanged. Divergence curves from correlated stride patterns exhibited delayed saturation compared to constant or random patterns (S2 Figures), providing evidence that fractal temporal structure prevents premature saturation of phase space exploration, thereby elevating long-term DE values. Consistent with this mechanism, three included studies [64,81,82] that applied stride-level time normalization by resampling each stride to a fixed number of data points (as in [52]) reported markedly suppressed long-term DE values ($\lambda L \approx 0.00$ in [81]; no significant condition effects in [64,82]).

Attention-dependent control mechanisms. The selective reduction of long-term DE during external pacing (Table 3) points to modifications in motor control strategy. Sensorimotor synchronization with rhythmic cues induces error correction processes that adjust each stride's timing to match external targets [34]. This correction mode contrasts with the loosely coupled, endogenous control of self-paced walking [83]. The corrections around target pace values generate oscillations at short timescales while suppressing the low-frequency variations that characterize normal gait [84]. In fractal analysis terms (DFA), short-term oscillations manifest as anti-persistent temporal structure, whereas low-frequency variations represent the persistent, long-range correlations that define high-complexity gait dynamics. Experimental evidence shows that indeed metronome walking induces anti-persistent stride interval dynamics [76,85,86]. This transformation from persistent to anti-persistent temporal structure constrains phase space exploration—trajectories converge more rapidly within the attractor—thereby accelerating divergence curve saturation and reducing long-term DE values. This is evident in divergence curves published in [40,50,57] (see S2 Figures). The reductions in both long-term DE and DFA scaling exponents during cued walking [47,51] thus reflect a common underlying mechanism: the alteration of fractal correlation structure through attention-dependent error correction mechanisms.

Dual constraint effects. The minimal sensitivity of long-term DE to constant-speed treadmill constraint (Table 3) reflects parameter-specific alterations in temporal correlation structure. Evidence from DFA demonstrates that treadmill walking selectively induces anti-persistent dynamics in stride-speed fluctuations while preserving persistent correlations in both stride intervals and stride lengths [76]. This selective transformation arises because maintaining constant speed requires coordinated adjustments between stride time and stride length—participants modulate both parameters in tandem to prevent positional drift, thereby

tightening speed control without imposing rigid constraints on either time series individually [34]. DEs computed from center-of-pressure trajectories or trunk acceleration signals likely capture a global fluctuation pattern, rather than the parameter-specific correlation structures revealed by DFA of individual time series. When only stride-speed correlations become anti-persistent while stride interval and stride length correlations remain persistent, this partial disruption produces moderate effects on the global attractor dynamics quantified by long-term DE. This interpretation is supported by the graded effects of constraint combinations observed in Table 3. Metronome synchronization during overground walking—which selectively modifies stride interval correlations while stride length and speed patterns remain persistent [85]—produces moderate long-term DE reductions (-26% to -40% in [47], -13% to -30% in [51], Table 3). In contrast, superimposing metronome pacing on treadmill walking generates larger reductions (−66% to −86%, Table 3), reflecting combined modification of both stride interval and stride length correlation structures. This graded response demonstrates that dual constraints reorganize attractor dynamics more profoundly than either single constraint alone, consistent with the parameter-specific nature of correlation structure alterations captured by DFA.

Long-term divergence as an index of attractor complexity. Taken together, this evidence suggests that long-term DE reflects gait complexity rather than stability: it correlates with fractal scaling exponents, responds to manipulation of stride interval correlation structure while short-term DE remains unchanged, and exhibits graded sensitivity to constraint combinations that selectively modify correlations in different gait parameters. These findings support reinterpreting long-term DE as an "Attractor Complexity Index" that captures how temporal correlation structure shapes phase space exploration.

Long-term divergence and gait automaticity

This section examines long-term divergence exponent patterns across two complementary lines of evidence: between-subject comparisons in clinical populations, and within-subject perturbation studies that experimentally manipulate gait demands.

Between-subject studies. Among the studies identified in this review, only eight met methodological criteria for valid between-subject comparisons (Table 2), having employed sufficiently long walking bouts and time-normalized signals to ensure reliable long-term divergence exponent estimation. The remaining clinical comparison studies, while potentially informative, warrant cautious interpretation. Within these methodologically rigorous studies, long-term DEs exhibited heterogeneous patterns across clinical populations. While short-term divergence exponents consistently increased in patients relative to controls, long-term

divergence exponents showed both reductions and increases, which requires specific analysis separated by conditions and pathologies.

Pain. Two included studies examining pain-related conditions converged on a consistent pattern: patients with chronic lower-limb pain [72] and those with knee osteoarthritis [75], demonstrated reduced long-term divergence. An earlier longitudinal study of knee osteoarthritis patients ([87], N=16) awaiting replacement surgery provides additional support for pain-related complexity reductions. Using only 30 strides per analysis, the study nevertheless used global time normalization and generalized estimating equations that appropriately modeled speed-dependent effects. At a clinically relevant speed of 3 km/h, the affected limb demonstrated a dissociation between stability and complexity measures: short-term divergence exponents were unchanged between patients and controls, while long-term divergence exponents showed 19% reduction.

Pain may induce a shift toward more controlled, less automatic gait characterized by greater attentional involvement in motor planning [36]. Pain-related attentional focus on bodily sensations and movement execution likely constrains the natural stride-to-stride variability that characterizes automatic walking.

Peripheral Neuropathy. Only one study met methodological criteria for valid between-subject comparison (Table 2). However, the foundational work on this topic comes from the two earlier studies by Dingwell and colleagues [41,88] examining the same cohort of diabetic neuropathy patients (N=26). While these studies used sufficiently long walking bouts (10 minutes continuous), they did not use time-normalized signals. Nevertheless, they observed a consistent pattern: patients with diabetic neuropathy showed reduced long-term divergence exponents compared to controls (-26%). The pattern observed by Manor et al. [71] using 100 consecutive strides with time-normalized signals confirmed this finding with a 4% reduction in long-term DE. In contrast, the within-subject study that examined acute plantar desensitization in healthy adults showed an opposite effect: both short-term and long-term DE increased ([65], Table 1). Such a result may point either to differential acute and habituation effects on gait dynamics or to heterogeneity in study design and measurement protocols.

Meta-analyses show that neuropathy patients walk more cautiously, with slower walking speed, shorter steps, and longer stance time [89,90]. Reduced plantar sensory information may make gait control more dependent on cognitive resources in diabetic patients compared to controls [90]. This higher cognitive demand is reflected in poorer dual-task performance [91]. Notably, these cautious gait adaptations are further amplified by cognitive impairment in diabetic patients, independent of neuropathy status [92]. The observed reduction in long-

term DEs may therefore reflect a compensatory shift toward a greater conscious control of gait characterized by a reduced reliance on automatic motor control.

Aging and fall risk. Three studies examining young versus older adults were excluded from this review due to methodological limitations. These studies generally showed trends toward increased short-term and long-term DE in older populations [93], though the long-term DE increases rarely reached statistical significance [94,95]. Among the four methodologically valid age-comparison studies (Table 2), three converged on reduced long-term DE in older adults [33,51,70], with reductions amplified under conditions that increase attentional demands such as visual perturbations or metronome cueing. Only Bizovska et al. [73] reported variable results, though their older group (mean age 57.5 years) represented middle-aged rather than elderly adults.

Regarding fall risk, two studies comparing older fallers to non-fallers were excluded from the between-subject comparison. Toebes et al. [46] (N=134, mean age 62.4 years) used non-time-normalized angular velocity signals in composite high-dimensional state spaces, while Amirpourabasi et al. [39] (N=34, mean age 69.3 years) used non-time-normalized joint kinematics and trunk acceleration. Both studies found that short-term rather than long-term DE discriminated fall status. Nevertheless, the principal component analysis by Amirpourabasi et al. retained long-term DE along anteroposterior and vertical axes, which were independently identified as the most sensitive to complexity changes [50], suggesting that long-term DE may carry fall-related information that their analytical framework did not fully exploit. The only included fall-risk study, a prospective one-year follow-up by Bizovska et al. [74] (N=131), found no clear long-term DE difference between non-fallers and single fallers, but a more consistent reduction in frequent fallers (−7% to −18%).

Age-related decline in gait automaticity is well-documented [25], with older adults demonstrating increased prefrontal cortex activation during usual walking compared to standing [96], though patterns of cortex recruitment during cognitively demanding walking may differ from younger adults [97]. This shift from automatic to more controlled gait regulation has been attributed to age-related changes in neural networks, where diminished function in cerebellum and basal ganglia regions necessitates greater compensatory engagement of attention-related cortical networks, particularly the prefrontal cortex [98–100]. This compensatory reliance on executive resources appears even more pronounced in older adults at risk of falling: fallers exhibit greater prefrontal cortex activation than non-fallers during cognitively demanding tasks [101], and prospective evidence indicates that higher prefrontal activation during dual-task walking independently predicts incident falls [102]. The observed reductions in long-term DE in older adults, particularly under conditions

945 that increase attentional demands such as visual perturbations, may therefore reflect this
946 transition toward less automatic gait control.
947 **Perturbation studies.** Regarding perturbation studies, six studies were excluded from
948 Table 1 due to methodological constraints. One study [64] used stride-level temporal
949 normalization, which suppresses long-term divergence [52] and precludes any interpretation
950 of long-term DE findings. Among the remaining excluded studies, two examined surface
951 perturbations with contrasting results. Chang et al. [20] found significantly reduced long-
952 term DE on a compliant foam surface compared to flat overground walking. In contrast,
953 Emmerzaal et al. [103] reported increased long-term DE across six non-flat outdoor surfaces
954 compared to flat-even walking. This discrepancy may partly reflect methodological limitations
955 (short walking bouts, no time normalization). However, it may also reflect a genuine
956 distinction between acute laboratory perturbations and familiar environmental terrain
957 variation, which may not necessarily disrupt automatic control.
958 Among the eight retained treadmill studies (Table 1), a remarkably consistent pattern
959 emerged: mechanical, visual, and vestibular perturbations increased short-term DE (+3% to
960 +86%) while simultaneously decreasing long-term DE (−4% to −51%), with the notable
961 exception of plantar desensitization which increased both measures. This is consistent with
962 the hypothesis that individuals allocate greater attentional resources to gait control to
963 manage the imposed challenges.
964 Neuroimaging studies provide direct support for this interpretation. Unpredictable slip
965 perturbations during treadmill walking induce significant increases in prefrontal cortex
966 activation as measured by functional near-infrared spectroscopy [104]. Visual perturbations
967 modulate electrocortical dynamics during treadmill walking, revealing increased cortical
968 processing demands under perturbed sensory conditions [105]. Last but not least, a
969 systematic review of mobile brain imaging studies further confirmed that unpredicted
970 mechanical perturbations elicit prominent cortical activation in frontal regions [106].
971 **Long-term divergence as an index of gait automaticity.** Across clinical
972 populations, experimental perturbations, and neuroimaging evidence, a coherent pattern
973 emerges: conditions that increase cortical involvement in gait control consistently reduce
974 long-term DE, while short-term DE tracks dynamic stability independently. Pain, neuropathy,
975 and aging each shift gait toward greater executive oversight, and this shift is mirrored by
976 reduced long-term divergence. Perturbation studies confirm this relationship experimentally,
977 with neuroimaging demonstrating that the same manipulations that decrease long-term DE
978 simultaneously increase prefrontal and frontal cortical activation. These converging lines of
979 evidence support the interpretation that long-term DE captures the degree of gait

automaticity, that is the extent to which locomotor control operates through automatic, subcortically-mediated processes rather than attention-dependent executive control.

Methodological considerations

Measurement methods. The included studies employed diverse measurement approaches (Fig 3). Despite this methodological heterogeneity, examination of response patterns across experimental manipulations reveals consistent findings, suggesting that long-term DE capture aspects of gait control that transcend specific measurement modalities. The critical prerequisite is likely that the measured signal captures the rhythmic pattern of locomotor cycles with their inherent stride-to-stride variations. Whether these dynamics are sampled through trunk accelerations, joint angle trajectories, or center-of-pressure displacements, the resulting measures converge on the same underlying phenomenon: the temporal organization of consecutive gait cycles.

Among the measurement approaches employed, trunk-level assessment offers particular advantages for capturing global gait coordination patterns. A majority of included studies (34/62, 55%) prioritized trunk measurements. Dingwell and Kang [78] emphasized that trunk movements reflect the integrated outcome of bilateral limb coordination and postural adjustments required to maintain upright progression. For assessments targeting the temporal organization of overall locomotor patterns, trunk-level measurements therefore represent the preferred approach, offering an integrated readout of stride-to-stride control dynamics. Similarly, center-of-pressure trajectories derived from instrumented treadmills (used in 3 included studies), although recorded at foot-ground interface level, represent another global measure of locomotor control: the center of pressure reflects the net ground reaction force application point and thus integrates all postural and propulsive adjustments required to regulate balance and forward progression during walking.

Measurement duration. Assessment of long-term DE as complexity measures imposes specific data duration requirements that differ from stability assessments. The basis for these requirements arises from the multiscale nature of gait complexity. Low-frequency variations, which are typical of persistent temporal correlations, manifest only across extended observation windows. Truncated recordings constrain the observable correlation timescales, filtering out the very low-frequency components. This explains why studies employing shorter recordings (fewer than 50 strides) may capture local stability properties adequately but potentially mischaracterize attractor complexity.

Methodological investigations within the included studies provide empirical guidance for minimum data requirements. Bruijn et al. [107] showed that divergence exponent precision improves with recording length, identifying approximately 150 consecutive strides as necessary for adequate measurement precision. Reynard and Terrier [61] confirmed these requirements, showing that 70-stride estimates yielded substantially higher reliability than 35-stride estimates for long-term DE. Riva et al. [108] observed that long-term divergence measures required a high stride count (110+ strides for 50% reliability threshold) compared to other stability metrics.

A 2020 systematic review on DFA methodology [109] established consensus guidelines recommending at least 500-600 consecutive strides for reliable estimation of scaling exponents, acknowledging that shorter recordings may inadequately capture the long-range correlation structure defining fractal gait dynamics. The more modest stride requirements for long-term DE (150-200 strides) compared to DFA's 500-600 stride recommendation suggest practical advantages for clinical gait assessment, requiring shorter recording durations while still capturing essential features of temporal correlation structure through phase space reconstruction [50]. Future investigations employing long-term divergence exponents as complexity measures should prioritize recording durations of at least 150-200 consecutive strides to ensure adequate precision for detecting meaningful differences while maintaining practical feasibility for clinical populations.

Data length effects and time normalization. Beyond precision considerations, measurement duration influences DE magnitudes. Bruijn et al. [107] demonstrated that both short-term and long-term DEs increase with data series length. As recording duration extends, the phase space trajectory explores increasingly distant regions of the attractor, and allows the algorithm to identify nearest neighbors separated by larger distances. These more distant neighbor pairs exhibit different divergence characteristics than the tightly clustered neighbors identified in shorter recordings. Reynard and Terrier confirmed this effect in a larger sample (N=95) [61]. Consequently, standardizing recording duration across experimental conditions and participant groups constitutes a critical methodological requirement, as divergence exponent differences attributable to unequal measurement lengths may confound or obscure true physiological differences between populations or conditions.

Study quality and risk of bias

Analytical rigor and methodological transparency. Quality assessment revealed good methodological transparency across the included studies (S3 Table), with comprehensive reporting of experimental protocols, measurement systems, signal processing procedures, and algorithm parameters necessary for study replication. This result may indicate that the field has achieved methodological maturity, with recent studies providing sufficient detail to enable independent replication. However, it is important to acknowledge that these favorable quality metrics partly reflect our inclusion criteria, which required studies to provide adequate methodological information to confirm use of Rosenstein's algorithm for long-term DE computation.

Outcome reporting: descriptive statistics. While analytical procedures were generally well-documented, the reporting of study outcomes showed considerable variability. A substantial proportion of studies, particularly those published before 2013, presented results exclusively through figures without accompanying numerical tables containing means and measures of variability. This reporting approach necessitated manual extraction of approximate values using specialized digitization software to enable quantitative synthesis and calculation of percentage changes between conditions or groups (as presented in Tables 2 and 3). The reliance on visual estimation from figures introduces measurement error and limits the precision of secondary analyses. Moreover, some studies reported only qualitative descriptions of differences or provided statistical test results without the underlying descriptive statistics, making meta-analytical synthesis unfeasible. A clear temporal improvement emerged after 2013, with recent studies increasingly providing complete descriptive statistics for all experimental conditions.

Outcome reporting: effect sizes and clinical significance. Beyond descriptive statistics, a more persistent limitation concerned the reporting and interpretation of effect sizes. Many studies across all time periods relied primarily on p-values and significance testing to characterize findings, with only 40% achieving full scores through comprehensive effect size reporting. Many studies reported results dichotomously as "significant" or "non-significant" without quantifying the magnitude of observed differences or discussing their potential clinical relevance. This approach limits the ability to distinguish between statistically significant but trivially small effects and meaningfully large effects that may warrant clinical attention. The absence of standardized effect size measures (such as Cohen's d, confidence intervals, or percentage differences with precision estimates) also constrains meta-analytic synthesis and prevents assessment of whether effects are consistent or variable across studies and populations. While some recent studies have begun incorporating effect size

reporting and discussing practical significance, this remains an underdeveloped aspect of reporting standards in the field.

Sample size considerations and recruitment practices. Domain 3 (sample size adequacy) achieved an average score of 1.7 out of 3 points, reflecting the modest sample sizes typical of exploratory biomechanical research (S3 Table). The median sample size across studies was 23 participants, with most studies enrolling fewer than 50 individuals. Given the exploratory and methodological nature of much of this research, these sample sizes are not inappropriate, particularly considering that most studies used within-subjects designs and repeated measures that enhance statistical power relative to between-subjects comparisons. However, only a minority of studies (<10%) provided explicit sample size justifications or reported formal power calculations, making it difficult to assess whether samples were adequate for detecting meaningful effects. A more substantive concern relates to recruitment practices and sample representativeness. Although rarely discussed explicitly in the reviewed articles, convenience sampling appeared widespread, with many studies likely recruiting from institutional populations such as university students or laboratory staff members. This practice may introduce systematic biases, as such samples tend to overrepresent younger, healthier, more educated, and potentially more physically active individuals compared to the general population. For studies comparing healthy controls to clinical populations, the use of convenience samples may exaggerate group differences if controls represent an unusually homogeneous and high-functioning subset of the healthy population. While these recruitment practices are understandable during the methodological development phase of a field, the transition toward clinical application requires more systematic attention to sample representativeness, demographic diversity, and explicit justification of sample sizes relative to the precision needed for clinical decision-making.

Limitations of the systematic review

Study identification and selection limitations. This systematic review has several limitations that should be acknowledged. First, the favorable quality assessment results partly reflect selection bias inherent to our inclusion criteria. Studies were required to provide sufficient methodological detail, resulting in exclusion of less transparently reported research during full-text screening. Second, while our database searches applied no language restrictions to ensure comprehensive literature retrieval, our protocol specified that full-text screening would be limited to English and French publications based on the linguistic competencies of our review team. The language restriction remained theoretical rather than

operational, as no relevant non-English/French studies required exclusion. Nevertheless, we acknowledge that future updates to this review would benefit from broader linguistic inclusion with appropriate translation resources to ensure maximum comprehensiveness. Third, our focus on peer-reviewed journal articles resulted in exclusion of grey literature including doctoral dissertations, conference abstracts, and technical reports, which may contain relevant methodological developments or findings not yet published in traditional venues. Fourth, publication bias may affect our synthesis, as studies reporting null findings or unsuccessful applications of the method may be less likely to appear in peer-reviewed literature. This bias may have been particularly pronounced given that most studies interpreted long-term DE as a stability measure: researchers observing paradoxical patterns—such as long-term DE reductions during destabilizing perturbations (Table 1)—may have regarded such findings as methodological failures rather than meaningful results, leading to selective non-reporting of studies with theoretically inconsistent outcomes under the traditional stability interpretation.

Synthesis limitations. The scope of quantitative synthesis was constrained by several factors. Despite identifying 62 relevant studies, comprehensive meta-analysis across all research questions was not feasible due to heterogeneous outcome reporting, particularly in earlier studies that presented results exclusively through figures without numerical tables. Only five studies provided the necessary correlation data for our meta-analysis examining associations between long-term DE and DFA scaling exponents, limiting the precision of this key finding. Similarly, only eight studies met methodological criteria for valid inclusion in between-subject clinical comparisons (Table 2), as most clinical studies used insufficiently long recording durations or lacked time normalization procedures necessary for reliable complexity assessment. The diverse measurement contexts—including different walking conditions, signal types, and body locations, while potentially strengthening the generalizability of findings, precluded direct quantitative comparison of effect magnitudes across studies. These synthesis limitations reflect the exploratory developmental stage of the field rather than fundamental obstacles to future evidence accumulation, as recent improvements in reporting standards suggest that subsequent systematic reviews will be better positioned for comprehensive meta-analytic synthesis.

Future directions and clinical translation

Research priorities and methodological advances. Several methodological priorities emerge from this review to advance the field toward clinical application. First, standardization of measurement protocols should prioritize recording durations of at least

150-200 consecutive strides for complexity assessment. Second, theoretical development remains incomplete. Establishment of clinically meaningful change thresholds for long-term DE is essential—what magnitude of reduction indicates functionally significant loss of gait automaticity? Development of age-stratified and population-specific normative reference values would enable interpretation of individual measurements relative to expected ranges. Third, validation of the complexity interpretation requires experimental studies that directly manipulate automaticity (e.g., dual-task paradigms, attention-demanding environments) while simultaneously measuring long-term DE and neuroimaging markers of cortical engagement. Finally, longitudinal research tracking long-term DE changes across disease progression, rehabilitation, or aging trajectories would clarify whether complexity measures predict clinically relevant outcomes such as fall risk, functional decline, or treatment response. Multi-site replication studies ensuring generalizability across diverse populations and measurement contexts remain particularly valuable given the current concentration of research within specialized laboratories.

Clinical translation and practical applications. Long-term DE holds promise as a complementary clinical measure that captures aspects of motor control not reflected in traditional gait metrics. Its potential clinical utility spans several domains. For diagnostic applications, reduced long-term DE may identify early-stage neurological changes characterized by diminished gait automaticity before gross functional impairments become apparent. In rehabilitation contexts, long-term DE could monitor recovery of automatic motor control following neurological injury or track treatment effects where restoration of natural gait patterns represents a therapeutic goal. For fall risk assessment, integration of complexity measures with conventional stability metrics may improve risk stratification. The method offers practical advantages over alternative complexity measures such as DFA: the more modest recording requirements (150-200 versus 500-600 strides) enhance feasibility for clinical populations with limited walking tolerance. Long-term DE should complement rather than replace existing clinical gait assessment tools, contributing one element within comprehensive evaluation protocols. With these foundations established, wearable sensor technologies could enable free-living complexity monitoring, extending assessment beyond laboratory settings to capture gait automaticity in natural environments where attentional demands and environmental challenges reflect real-world function.

Conclusion

This systematic review provides converging evidence that long-term DE measures gait complexity and automaticity rather than stability. The substantial correlation with DFA scaling exponents (r = 0.64), combined with systematic dissociation from short-term stability measures during perturbations and cueing, supports reinterpreting this metric as the Attractor Complexity Index—a marker of the temporal organization and attentional demands of locomotor control.

This paradigm shift has practical implications. Rather than duplicating information from short-term stability measures, long-term divergence may uniquely capture motor-cognitive aspects of gait control relevant to fall risk, disease progression, and rehabilitation effectiveness. Establishing this measure as a complementary biomarker for gait automaticity would enhance multidimensional assessment of locomotor function in clinical populations.

Acknowledgments

We thank Marco Pedrotti and Florence Waelchli for their assistance with administrative tasks.

REFERENCES

1. Perry J. Gait Analysis: Normal and Pathological Function, Second Edition. 2nd ed. Thorofare: SLACK, Incorporated; 2010.

2. Singh SJ, Puhan MA, Andrianopoulos V, Hernandes NA, Mitchell KE, Hill CJ, et al. An official systematic review of the European Respiratory Society/American Thoracic Society: measurement properties of field walking tests in chronic respiratory disease. Eur Respir J. 2014;44: 1447–1478. doi:10.1183/09031936.00150414

3. Giannitsi S, Bougiakli M, Bechlioulis A, Kotsia A, Michalis LK, Naka KK. 6-minute walking test: a useful tool in the management of heart failure patients. Ther Adv Cardiovasc Dis. 2019;13: 1753944719870084. doi:10.1177/1753944719870084

4. Yogev G, Hausdorff JM, Giladi N. The Role of Executive Function and Attention in Gait. Mov Disord. 2008;23: 329–472. doi:10.1002/mds.21720

5. Perera S, Patel KV, Rosano C, Rubin SM, Satterfield S, Harris T, et al. Gait Speed Predicts Incident Disability: A Pooled Analysis. The Journals of Gerontology: Series A. 2016;71: 63–71. doi:10.1093/gerona/glv126

6. Montero-Odasso M, Verghese J, Beauchet O, Hausdorff JM. Gait and Cognition: A Complementary Approach to Understanding Brain Function and the Risk of Falling. Journal of the American Geriatrics Society. 2012;60: 2127–2136. doi:10.1111/j.1532-5415.2012.04209.x

7. Allali G, Launay CP, Blumen HM, Callisaya ML, De Cock A-M, Kressig RW, et al. Falls, Cognitive Impairment, and Gait Performance: Results From the GOOD Initiative. J Am Med Dir Assoc. 2017;18: 335–340. doi:10.1016/j.jamda.2016.10.008

8. Verghese J, Wang C, Lipton RB, Holtzer R, Xue X. Quantitative gait dysfunction and risk of cognitive decline and dementia. Journal of Neurology, Neurosurgery & Psychiatry. 2007;78: 929–935. doi:10.1136/jnnp.2006.106914

9. Dommershuijsen LJ, Isik BM, Darweesh SKL, van der Geest JN, Ikram MK, Ikram MA. Unraveling the Association Between Gait and Mortality-One Step at a Time. J Gerontol A Biol Sci Med Sci. 2020;75: 1184–1190. doi:10.1093/gerona/glz282

10. Tinetti ME. Performance-Oriented Assessment of Mobility Problems in Elderly Patients. Journal of the American Geriatrics Society. 1986;34: 119–126. doi:10.1111/j.1532-5415.1986.tb05480.x

11. Hulleck AA, Menoth Mohan D, Abdallah N, El Rich M, Khalaf K. Present and future of gait assessment in clinical practice: Towards the application of novel trends and technologies. Front Med Technol. 2022;4. doi:10.3389/fmedt.2022.901331

12. Sprott JC. Chaos and time-series analysis. Oxford ; New York: Oxford University Press; 2003.

13. Stergiou N, Harbourne R, Cavanaugh J. Optimal movement variability: a new theoretical perspective for neurologic physical therapy. J Neurol Phys Ther. 2006;30: 120–129. doi:10.1097/01.npt.0000281949.48193.d9

14. Moghtadaei M, Dorey TW, Rose RA. Evaluation of non-linear heart rate variability using multi-scale multi-fractal detrended fluctuation analysis in mice: Roles of the autonomic nervous system and sinoatrial node. Front Physiol. 2022;13. doi:10.3389/fphys.2022.970393

15. Dingwell JB. Lyapunov Exponents. Wiley Encyclopedia of Biomedical Engineering. John Wiley & Sons, Ltd; 2006. doi:10.1002/9780471740360.ebs0702

16. Hausdorff JM, Mitchell SL, Firtion R, Peng CK, Cudkowicz ME, Wei JY, et al. Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease. J Appl Physiol. 1997;82: 262–269. doi:10.1152/jappl.1997.82.1.262

17. Yentes JM, Raffalt PC. Entropy analysis in gait research – methodological considerations and recommendations. Ann Biomed Eng. 2021;49: 979–990. doi:10.1007/s10439-020-02616-8

18. Lukšys D, Griškevičius J. Application of nonlinear analysis for the assessment of gait in patients with Parkinson's disease. Technol Health Care. 2022;30: 201–208. doi:10.3233/THC-219003

19. Lahmiri S. Gait Nonlinear Patterns Related to Parkinson's Disease and Age. IEEE Transactions on Instrumentation and Measurement. 2019;68: 2545–2551. doi:10.1109/TIM.2018.2866316

20. Amirpourabasi A, Lamb SE, Chow JY, Williams GKR. Nonlinear Dynamic Measures of Walking in Healthy Older Adults: A Systematic Scoping Review. Sensors. 2022;22: 4408. doi:10.3390/s22124408

21. Freitas M, Pinho F, Pinho L, Silva S, Figueira V, Vilas-Boas JP, et al. Biomechanical Assessment Methods Used in Chronic Stroke: A Scoping Review of Non-Linear Approaches. Sensors. 2024;24: 2338. doi:10.3390/s24072338

22. Stergiou N, editor. Nonlinear analysis for human movement variability. Boca Raton London New York: CRC Press; 2016.

23. Zanin M, Olivares F, Pulido-Valdeolivas I, Rausell E, Gomez-Andres D. Gait analysis under the lens of statistical physics. Computational and Structural Biotechnology Journal. 2022;20: 3257–3267. doi:10.1016/j.csbj.2022.06.022

24. Stuss DT, Levine B. Adult clinical neuropsychology: lessons from studies of the frontal lobes. Annu Rev Psychol. 2002;53: 401–433. doi:10.1146/annurev.psych.53.100901.135220

25. Clark DJ. Automaticity of walking: functional significance, mechanisms, measurement and rehabilitation strategies. Front Hum Neurosci. 2015;9: 246. doi:10.3389/fnhum.2015.00246

26. Crowley P, Vuillerme N, Samani A, Madeleine P. The effects of walking speed and mobile phone use on the walking dynamics of young adults. Sci Rep. 2021;11: 1237. doi:10.1038/s41598-020-79584-5

27. Walsh GS. Dynamics of Modular Neuromotor Control of Walking and Running during Single and Dual Task Conditions. Neuroscience. 2021;465: 1–10. doi:10.1016/j.neuroscience.2021.04.004

28. Ahmadi S, Wu C, Sepehri N, Kantikar A, Nankar M, Szturm T. The Effects of Aging and Dual Tasking on Human Gait Complexity During Treadmill Walking: A Comparative Study Using Quantized Dynamical Entropy and Sample Entropy. J Biomech Eng. 2018;140. doi:10.1115/1.4037945

29. Lamoth CJ, van Deudekom FJ, van Campen JP, Appels BA, de Vries OJ, Pijnappels M. Gait stability and variability measures show effects of impaired cognition and dual tasking in frail people. J Neuroeng Rehabil. 2011;8: 2. doi:10.1186/1743-0003-8-2

30. Tajali S, Mehravar M, Negahban H, van Dieën JH, Shaterzadeh-Yazdi M-J, Mofateh R. Impaired local dynamic stability during treadmill walking predicts future falls in patients with multiple sclerosis: A prospective cohort study. Clin Biomech (Bristol). 2019;67: 197–201. doi:10.1016/j.clinbiomech.2019.05.013

31. Hove MJ, Suzuki K, Uchitomi H, Orimo S, Miyake Y. Interactive rhythmic auditory stimulation reinstates natural 1/f timing in gait of Parkinson's patients. PLoS One. 2012;7: e32600. doi:10.1371/journal.pone.0032600

32. Schaefer RS. Auditory rhythmic cueing in movement rehabilitation: findings and possible mechanisms. Philos Trans R Soc Lond B Biol Sci. 2014;369: 20130402. doi:10.1098/rstb.2013.0402

33. Franz JR, Francis CA, Allen MS, O'Connor SM, Thelen DG. Advanced age brings a greater reliance on visual feedback to maintain balance during walking. Human Movement Science. 2015;40: 381–392. doi:10.1016/j.humov.2015.01.012

34. Roerdink M, de Jonge CP, Smid LM, Daffertshofer A. Tightening Up the Control of Treadmill Walking: Effects of Maneuverability Range and Acoustic Pacing on Stride-to-Stride Fluctuations. Front Physiol. 2019;10. doi:10.3389/fphys.2019.00257

35. Decker LM, Cignetti F, Stergiou N. Executive function orchestrates regulation of task-relevant gait fluctuations. Gait Posture. 2013;38: 537–540. doi:10.1016/j.gaitpost.2012.12.018

36. Homs AF, Dupeyron A, Torre K. Relationship between gait complexity and pain attention in chronic low back pain. Pain. 2022;163: e31–e39. doi:10.1097/j.pain.0000000000002303

37. Toebes MJP, Hoozemans MJM, Furrer R, Dekker J, van Dieën JH. Associations between measures of gait stability, leg strength and fear of falling. Gait Posture. 2015;41: 76–80. doi:10.1016/j.gaitpost.2014.08.015

38. Bruijn SM, Meijer OG, Beek PJ, van Dieën JH. Assessing the stability of human locomotion: a review of current measures. J R Soc Interface. 2013;10: 20120999. doi:10.1098/rsif.2012.0999

39. Amirpourabasi A, Lamb SE, Chow JY, Williams GKR. Using nonlinear dynamic analysis to differentiate fall status in older women. GAIT & POSTURE. 2026;124. doi:10.1016/j.gaitpost.2025.110032

40. Terrier P, Dériaz O. Non-linear dynamics of human locomotion: effects of rhythmic auditory cueing on local dynamic stability. Front Physiol. 2013;4: 230. doi:10.3389/fphys.2013.00230

41. Dingwell JB, Cusumano JP, Sternad D, Cavanagh PR. Slower speeds in patients with diabetic neuropathy lead to improved local dynamic stability of continuous overground walking. Journal of Biomechanics. 2000;33: 1269–1277. doi:10.1016/S0021-9290(00)00092-0

42. Roos PE, Dingwell JB. Influence of Simulated Neuromuscular Noise on the Dynamic Stability and Fall Risk of a 3D Dynamic Walking Model. J Biomech. 2011;44: 1514–1520. doi:10.1016/j.jbiomech.2011.03.003

43. Bruijn SM, Bregman DJJ, Meijer OG, Beek PJ, van Dieën JH. Maximum Lyapunov exponents as predictors of global gait stability: a modelling approach. Med Eng Phys. 2012;34: 428–436. doi:10.1016/j.medengphy.2011.07.024

44. Mehdizadeh S. The largest Lyapunov exponent of gait in young and elderly individuals: A systematic review. Gait Posture. 2018;60: 241–250. doi:10.1016/j.gaitpost.2017.12.016

45. Lorenz EN. The Essence of Chaos. Nachdr. Seattle: Univ. of Washington Press; 2008.

46. Toebes MJP, Hoozemans MJM, Furrer R, Dekker J, Van Dieën JH. Local dynamic stability and variability of gait are associated with fall history in elderly subjects. Gait & Posture. 2012;36: 527–531. doi:10.1016/j.gaitpost.2012.05.016

47. Sejdić E, Fu Y, Pak A, Fairley JA, Chau T. The Effects of Rhythmic Sensory Cues on the Temporal Dynamics of Human Gait. Ernst MO, editor. PLoS ONE. 2012;7: e43104. doi:10.1371/journal.pone.0043104

48. Van Schooten KS, Sloot LH, Bruijn SM, Kingma H, Meijer OG, Pijnappels M, et al. Sensitivity of trunk variability and stability measures to balance impairments induced by galvanic vestibular stimulation during gait. Gait & Posture. 2011;33: 656–660. doi:10.1016/j.gaitpost.2011.02.017

49. Jordan K, Challis JH, Cusumano JP, Newell KM. Stability and the time-dependent structure of gait variability in walking and running. Human Movement Science. 2009;28: 113–128. doi:10.1016/j.humov.2008.09.001

50. Piergiovanni, S, Terrier, P. Effects of metronome walking on long-term attractor divergence and correlation structure of gait: a validation study in older people. Scientific Reports. 2024;14: 15784. doi:10.1038/s41598-024-65662-5

51. Piergiovanni S, Terrier P. Validity of Linear and Nonlinear Measures of Gait Variability to Characterize Aging Gait with a Single Lower Back Accelerometer. SENSORS. 2024;24. doi:10.3390/s24237427

52. Terrier P, Reynard F. Maximum Lyapunov exponent revisited: Long-term attractor divergence of gait dynamics is highly sensitive to the noise structure of stride intervals. Gait & Posture. 2018;66: 236–241. doi:10.1016/j.gaitpost.2018.08.010

53. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372: n71. doi:10.1136/bmj.n71

54. Schiavo JH. PROSPERO: An International Register of Systematic Review Protocols. Med Ref Serv Q. 2019;38: 171–180. doi:10.1080/02763869.2019.1588072

55. Terrier P. From stability to complexity: A systematic review protocol on long-term divergence exponents in gait analysis. medRxiv. 2024; 2024–12.

56. Terrier P, Dériaz O. Kinematic variability, fractal dynamics and local dynamic stability of treadmill walking. J NeuroEngineering Rehabil. 2011;8: 12. doi:10.1186/1743-0003-8-12

57. Terrier P. Complexity of human walking: the attractor complexity index is sensitive to gait synchronization with visual and auditory cues. PeerJ. 2019;7: e7417. doi:10.7717/peerj.7417

58. Pustejovsky JE, Tipton E. Meta-analysis with Robust Variance Estimation: Expanding the Range of Working Models. Prev Sci. 2022;23: 425–438. doi:10.1007/s11121-021-01246-3

59. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. Journal of Statistical Software. 2010;36: 1–48. doi:10.18637/jss.v036.i03

60. Pustejovsky JE. clubSandwich. 2023. doi:10.32614/CRAN.package.clubSandwich

61. Reynard P, Terrier P. Local dynamic stability of treadmill walking: Intrasession and week-to-week repeatability. Journal of Biomechanics. 2014;47: 74–80. doi:10.1016/j.jbiomech.2013.10.011

62. Takens F. Detecting strange attractors in turbulence. In: Rand D, Young L-S, editors. Dynamical Systems and Turbulence, Warwick 1980. Springer Berlin Heidelberg; 1981. pp. 366–381.

63. Dingwell JB, Cusumano JP, Cavanagh PR, Sternad D. Local Dynamic Stability Versus Kinematic Variability of Continuous Overground and Treadmill Walking. Journal of Biomechanical Engineering. 2001;123: 27–32. doi:10.1115/1.1336798

64. Yamagata S, Yamaguchi T, Shinya M, Milosevic M, Masani K. Comparison of sensitivity among dynamic balance measures during walking with different tasks. ROYAL SOCIETY OPEN SCIENCE. 2024;11. doi:10.1098/rsos.230883

65. Manor B, Wolenski P, Guevaro A, Li L. Differential effects of plantar desensitization on locomotion dynamics. Journal of electromyography and kinesiology : official journal of the International Society of Electrophysiological Kinesiology. 2009;19: e320-8. doi:10.1016/j.jelekin.2008.06.006

66. McAndrew PM, Wilken JM, Dingwell JB. Dynamic stability of human walking in visually and mechanically destabilizing environments. J Biomech. 2011;44: 644–649. doi:10.1016/j.jbiomech.2010.11.007

67. Sinitksi EH, Terry K, Wilken JM, Dingwell JB. Effects of perturbation magnitude on dynamic stability when walking in destabilizing environments. Journal of Biomechanics. 2012;45: 2084–2091. doi:10.1016/j.jbiomech.2012.05.039

68. Rábago CA, Dingwell JB, Wilken JM. Reliability and Minimum Detectable Change of Temporal-Spatial, Kinematic, and Dynamic Stability Measures during Perturbed Gait. PLoS ONE. 2015;10: e0142083. doi:10.1371/journal.pone.0142083

69. Reynard F, Terrier P. Role of visual input in the control of dynamic balance: variability and instability of gait in treadmill walking while blindfolded. Exp Brain Res. 2015;233: 1031–1040. doi:10.1007/s00221-014-4177-5

70. Qiao M, Truong KN, Franz JR. Does local dynamic stability during unperturbed walking predict the response to balance perturbations? An examination across age and falls history. Gait & Posture. 2018;62: 80–85. doi:10.1016/j.gaitpost.2018.03.011

71. Manor B, Wolenski P, Li L. Faster walking speeds increase local instability among people with peripheral neuropathy. Journal of Biomechanics. 2008;41: 2787–2792. doi:10.1016/j.jbiomech.2008.07.006

72. Terrier P, Le Carre J, Connaissa M-L, Leger B, Luthi F. Monitoring of Gait Quality in Patients With Chronic Pain of Lower Limbs. IEEE Trans Neural Syst Rehabil Eng. 2017;25: 1843–1852. doi:10.1109/TNSRE.2017.2688485

73. Bizovska L, Svoboda Z, Kubonova E, Vuillerme N, Hirjakova Z, Janura M. The differences between overground and treadmill walking in nonlinear, entropy-based and

frequency variables derived from accelerometers in young and older women - preliminary report. Acta of Bioengineering and Biomechanics. 2018;20: 93–100. doi:10.5277/ABB-00987-2017-02

74. Bizovska L, Svoboda Z, Janura M, Bisi MC, Vuillerme N. Local dynamic stability during gait for predicting falls in elderly people: A one-year prospective study. Plos One. 2018;13: e0197091. doi:10.1371/journal.pone.0197091

75. Matinazad N, Esteki A, Ghomashchi H. Comprehensive characterization of gait variability in patients with knee osteoarthritis for altered velocities. J Mech Med Biol. 2018;18: 1850041. doi:10.1142/S0219519418500410

76. Terrier P. Fractal Fluctuations in Human Walking: Comparison Between Auditory and Visually Guided Stepping. Ann Biomed Eng. 2016;44: 2785–2793. doi:10.1007/s10439-016-1573-y

77. Rosenstein MT, Collins JJ, De Luca CJ. A practical method for calculating largest Lyapunov exponents from small data sets. Physica D: Nonlinear Phenomena. 1993;65: 117–134. doi:10.1016/0167-2789(93)90009-P

78. Dingwell JB, Kang HG. Differences Between Local and Orbital Dynamic Stability During Human Walking. Journal of Biomechanical Engineering. 2007;129: 586–593. doi:10.1115/1.2746383

79. Ahn J, Hogan N. Long-range correlations in stride intervals may emerge from non-chaotic walking dynamics. PLoS One. 2013;8: e73239. doi:10.1371/journal.pone.0073239

80. Raffalt PC, Kent JA, Wurdeman SR, Stergiou N. To walk or to run – a question of movement attractor stability. J Exp Biol. 2020;223: jeb224113. doi:10.1242/jeb.224113

81. Rantalainen T, Hart NH, Nimphius S, Wundersitz DW. Associations Between Step Duration Variability and Inertial Measurement Unit Derived Gait Characteristics. Journal of Applied Biomechanics. 2016;32: 401–406. doi:10.1123/jab.2015-0266

82. Sedighi A, Rashedi E, Nussbaum MA. A head-worn display ("smart glasses") has adverse impacts on the dynamics of lateral position control during gait. Gait & Posture. 2020;81: 126–130. doi:10.1016/j.gaitpost.2020.07.014

83. Dingwell JB, Cusumano JP. Re-interpreting detrended fluctuation analyses of stride-to-stride variability in human walking. Gait Posture. 2010;32: 348–353. doi:10.1016/j.gaitpost.2010.06.004

84. Terrier P. Step-to-step variability in treadmill walking: influence of rhythmic auditory cueing. PLoS ONE. 2012;7: e47171. doi:10.1371/journal.pone.0047171

85. Terrier P, Turner V, Schutz Y. GPS analysis of human locomotion: further evidence for long-range correlations in stride-to-stride fluctuations of gait parameters. Human movement science. 2005;24: 97–115.

86. Terrier P, Dériaz O. Persistent and anti-persistent pattern in stride-to-stride variability of treadmill walking: influence of rhythmic auditory cueing. Human movement science. 2012;31: 1585–1597.

87. Yakhdani HR, Bafghi HA, Meijer OG, Bruijn SM, Dikkenberg NVD, Stibbe AB, et al. Stability and variability of knee kinematics during gait in knee osteoarthritis before and after replacement surgery. Clinical Biomechanics. 2010;25: 230–236. doi:10.1016/j.clinbiomech.2009.12.003

88. Dingwell JB, Cusumano JP. Nonlinear time series analysis of normal and pathological human walking. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2000;10: 848–863. doi:10.1063/1.1324008

89. Wang Z, Peng S, Zhang H, Sun H, Hu J. Gait Parameters and Peripheral Neuropathy in Patients With Diabetes: A Meta-Analysis. Front Endocrinol (Lausanne). 2022;13: 891356. doi:10.3389/fendo.2022.891356

90. Mustapa A, Justine M, Mohd Mustafah N, Jamil N, Manaf H. Postural Control and Gait Performance in the Diabetic Peripheral Neuropathy: A Systematic Review. Biomed Res Int. 2016;2016: 9305025. doi:10.1155/2016/9305025

91. Paul L, Ellis BM, Leese GP, McFadyen AK, McMurray B. The effect of a cognitive or motor task on gait parameters of diabetic patients, with and without neuropathy. Diabet Med. 2009;26: 234–239. doi:10.1111/j.1464-5491.2008.02655.x

92. Roman de Mettelinge T, Delbaere K, Calders P, Gysel T, Van Den Noortgate N, Cambier D. The impact of peripheral neuropathy and cognitive decrements on gait in older adults with type 2 diabetes mellitus. Arch Phys Med Rehabil. 2013;94: 1074–1079. doi:10.1016/j.apmr.2013.01.018

93. Cignetti F, Decker LM, Stergiou N. Sensitivity of the Wolf's and Rosenstein's Algorithms to Evaluate Local Dynamic Stability from Small Gait Data Sets. Ann Biomed Eng. 2012;40: 1122–1130. doi:10.1007/s10439-011-0474-3

94. Kang HG, Dingwell JB. Effects of walking speed, strength and range of motion on gait stability in healthy older adults. Journal of Biomechanics. 2008;41: 2899–2905. doi:10.1016/j.jbiomech.2008.08.002

95. Kang HG, Dingwell JB. Dynamic stability of superior vs. inferior segments during walking in young and older adults. Gait & Posture. 2009;30: 260–263. doi:10.1016/j.gaitpost.2009.05.003

96. Mirelman A, Maidan I, Bernad-Elazari H, Shustack S, Giladi N, Hausdorff JM. Effects of aging on prefrontal brain activation during challenging walking conditions. Brain and Cognition. 2017;115: 41–46. doi:10.1016/j.bandc.2017.04.002

97. Holtzer R, Mahoney JR, Izzetoglu M, Izzetoglu K, Onaral B, Verghese J. fNIRS Study of Walking and Walking While Talking in Young and Old Individuals. J Gerontol A Biol Sci Med Sci. 2011;66A: 879–887. doi:10.1093/gerona/glr068

98. Holtzer R, Epstein N, Mahoney JR, Izzetoglu M, Blumen HM. Neuroimaging of Mobility in Aging: A Targeted Review. J Gerontol A Biol Sci Med Sci. 2014;69: 1375–1388. doi:10.1093/gerona/glu052

99. Rosano C, Aizenstein HJ, Studenski S, Newman AB. A Regions-of-Interest Volumetric Analysis of Mobility Limitations in Community-Dwelling Older Adults. J Gerontol A Biol Sci Med Sci. 2007;62: 1048–1055. doi:10.1093/gerona/62.9.1048

100. Zwergal A, Linn J, Xiong G, Brandt T, Strupp M, Jahn K. Aging of human supraspinal locomotor and postural control in fMRI. Neurobiology of Aging. 2012;33: 1073–1084. doi:10.1016/j.neurobiolaging.2010.09.022

101. Halliday DrewWR, Hundza SR, Garcia-Barrera MA, Klimstra M, Commandeur D, Lukyn TV, et al. Comparing Executive Function, Evoked Haemodynamic Response, and Gait as Predictors of Variations in Mobility for Older Adults. J Clin Exp Neuropsychol. 2018;40: 151–160. doi:10.1080/13803395.2017.1325453

102. Verghese J, Wang C, Ayers E, Izzetoglu M, Holtzer R. Brain activation in high-functioning older adults and falls: Prospective cohort study. Neurology. 2017;88: 191–197. doi:10.1212/WNL.0000000000003421

103. Emmerzaal J, Ippersiel P, Dixon PC. Non-Linear Gait Dynamics Are Affected by Commonly Occurring Outdoor Surfaces and Sex in Healthy Adults. SENSORS. 2025;25. doi:10.3390/s25134191

104. Lee B-C, Choi J, Martin BJ. Roles of the prefrontal cortex in learning to time the onset of pre-existing motor programs. PLOS ONE. 2020;15: e0241562. doi:10.1371/journal.pone.0241562

105. Malcolm BR, Foxe JJ, Butler JS, Molholm S, De Sanctis P. Cognitive load reduces the effects of optic flow on gait and electrocortical dynamics during treadmill walking. Journal of Neurophysiology. 2018;120: 2246–2259. doi:10.1152/jn.00079.2018

106. Purohit R, Bhatt T. Mobile Brain Imaging to Examine Task-Related Cortical Correlates of Reactive Balance: A Systematic Review. Brain Sciences. 2022;12: 1487. doi:10.3390/brainsci12111487

107. Bruijn SM, Van Dieën JH, Meijer OG, Beek PJ. Statistical precision and sensitivity of measures of dynamic gait stability. Journal of Neuroscience Methods. 2009;178: 327–333. doi:10.1016/j.jneumeth.2008.12.015

108. Riva F, Bisi MC, Stagni R. Gait variability and stability measures: Minimum number of strides and within-session reliability. Computers in Biology and Medicine. 2014;50: 9–13. doi:10.1016/j.compbiomed.2014.04.001

109. Ravi DK, Marmelat V, Taylor WR, Newell KM, Stergiou N, Singh NB. Assessing the Temporal Organization of Walking Variability: A Systematic Review and Consensus Guidelines on Detrended Fluctuation Analysis. Front Physiol. 2020;11: 562. doi:10.3389/fphys.2020.00562

Supporting Information

S1 Text. Systematic review database. Complete data extraction tables with study characteristics (Table A: research types; Table B: study characteristics; Table C:

methodological specifications), search strategy, and detailed information for all 62 included studies. (PDF)

S2 Figures. Divergence curve figures. Compiled divergence curves from 30 studies with detailed figure descriptions showing methodological illustrations and comparative results. (PDF)

S3 Table. Quality assessment results. Complete quality assessment scores for all 62 studies using the three-domain framework (maximum 15 points). (XLSX)

Figure legends

Fig 1. PRISMA flowchart of systematic review study selection. Flow diagram showing identification, screening, eligibility assessment, and final inclusion of studies examining long-term divergence exponents in human gait analysis.

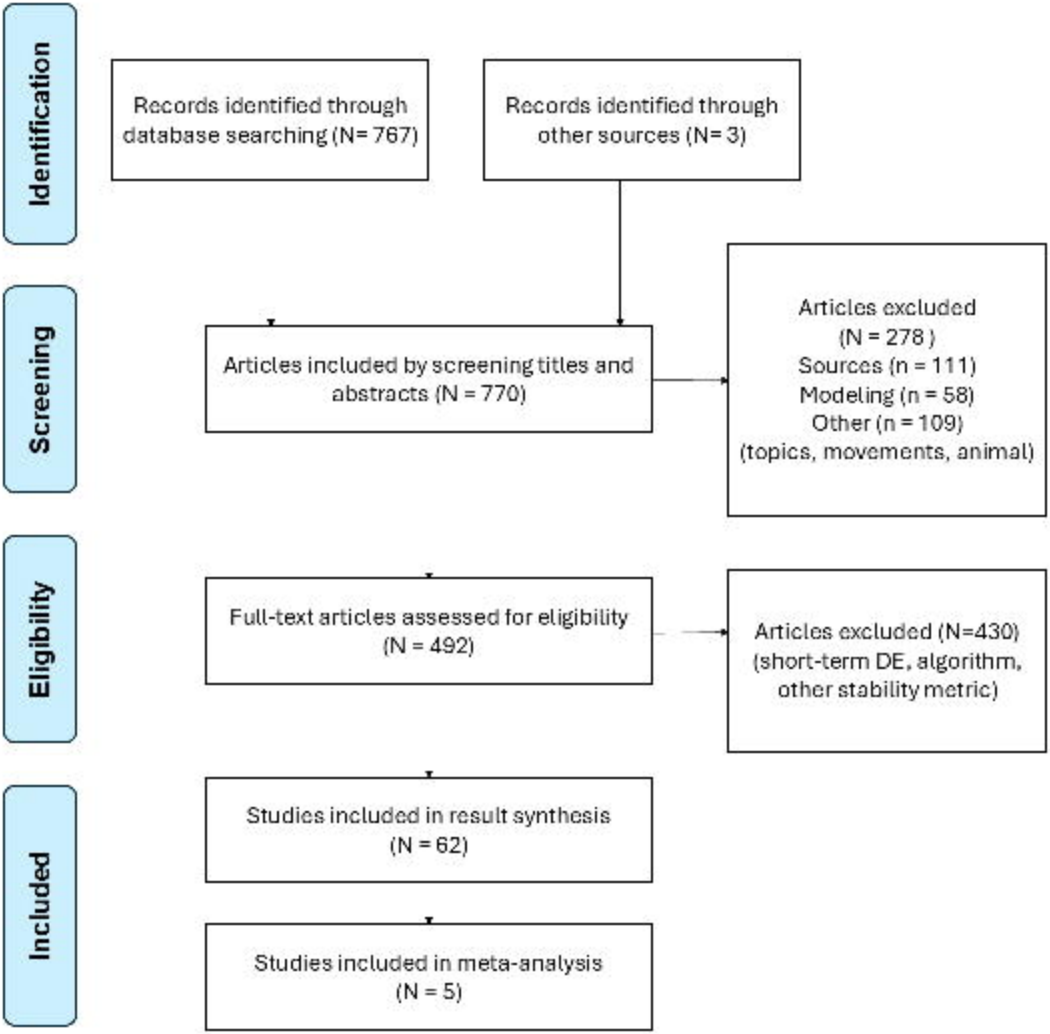
Fig 2. Study characteristics. a) Annual publication trends. b) Distribution by research type and participant category. c) Sample size distribution (n = 63 because one study reported two independent experiments with separate participant groups). DE, divergence exponent.

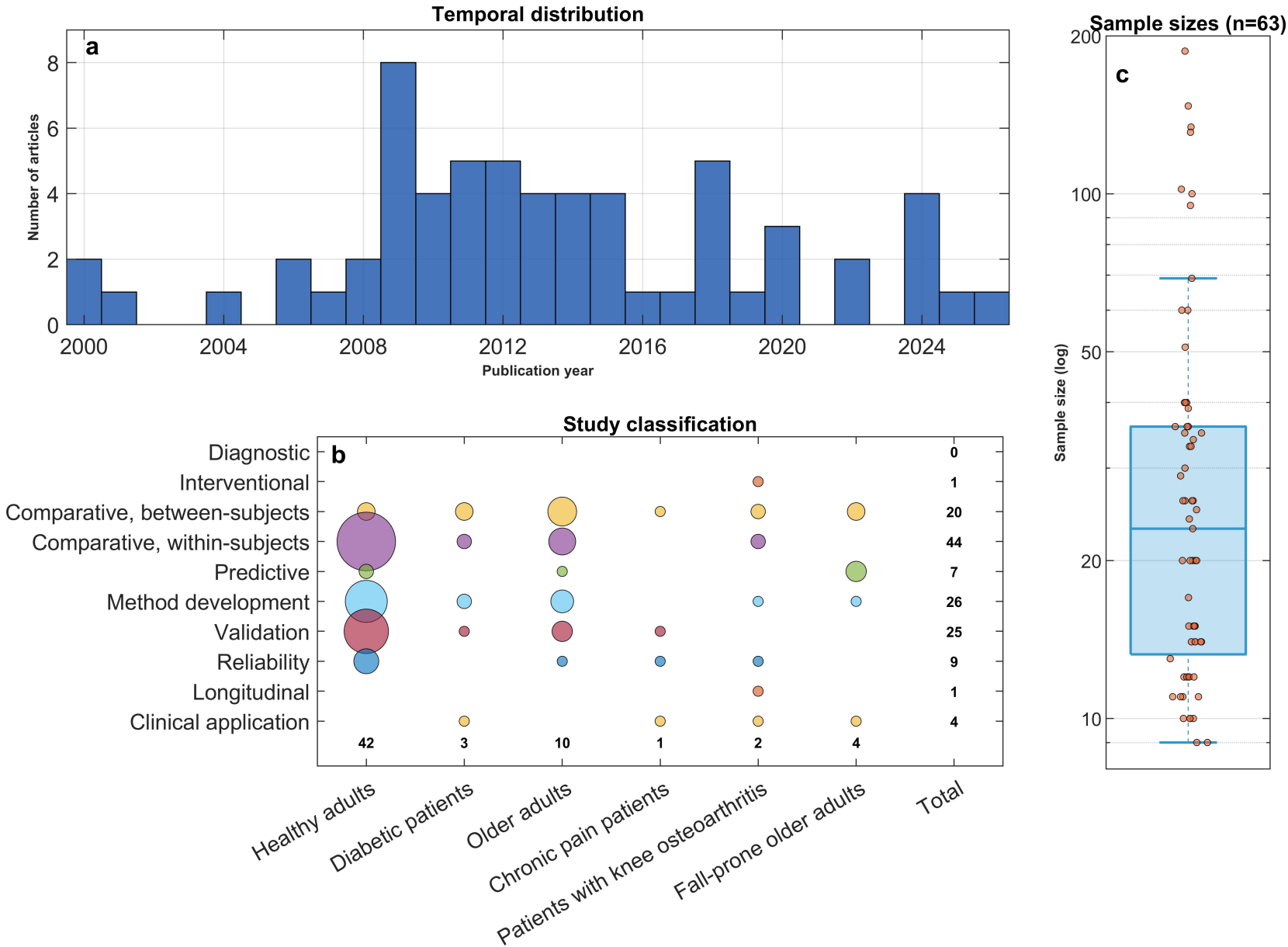
Fig 3. Methodological characteristics. a) Walking conditions. **b)** Measurement systems. **c)** Anatomical measurement sites. IMU, inertial measurement unit.

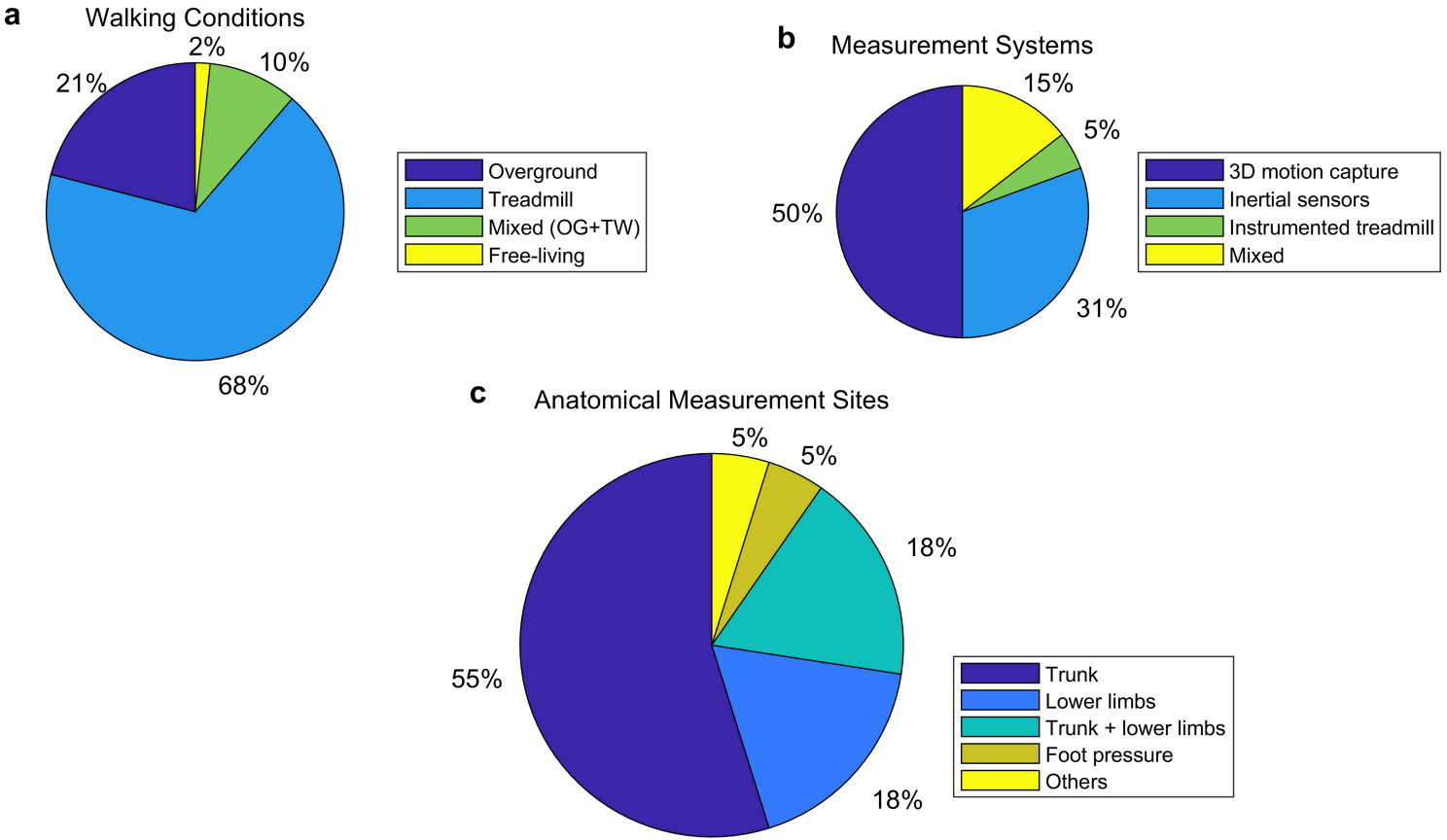
Fig 4. Measurement duration. Distribution of walking bout durations across four duration categories.

Fig 5. Signal processing characteristics. a) Signal sources used for computing DEs. **b)** Resampling methods to normalize signal durations. DE, divergence exponent; COP, center of pressure; EMG, electromyography.

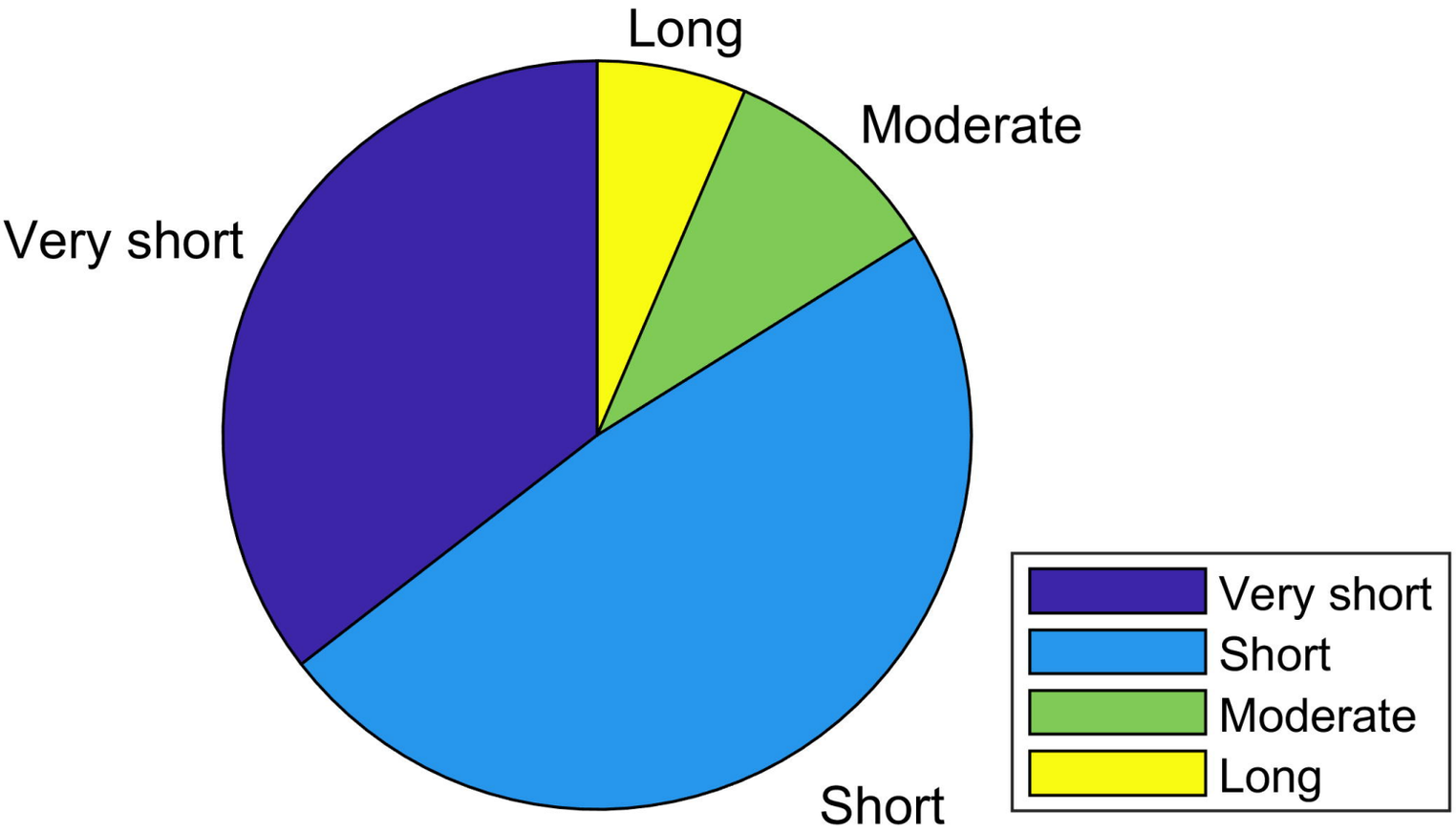
Fig 6. Meta-analysis of correlations between long-term DE and DFA scaling exponents. Forest plot showing study-level correlations with 95% confidence intervals and pooled estimate using correlated and hierarchical effects (CHE) model with robust variance estimation. N, total observations; k, number of correlation estimates per study; CI, confidence interval; DFA, detrended fluctuation analysis; DE, divergence exponent.


Identification
Records identified through database searching (N= 767)
Records identified through other sources (N= 3)
Screening
Articles included by screening titles and abstracts (N = 770)
Articles excluded
(N = 278)
Sources (n = 111)
Modeling (n = 58)
Other (n = 109)
(topics, movements, animal)
Eligibility
Full-text articles assessed for eligibility
(N = 492)
Articles excluded (N=430)
(short-term DE, algorithm, other stability metric)
Included
Studies included in result synthesis
(N = 62)
Studies included in meta-analysis
(N = 5)

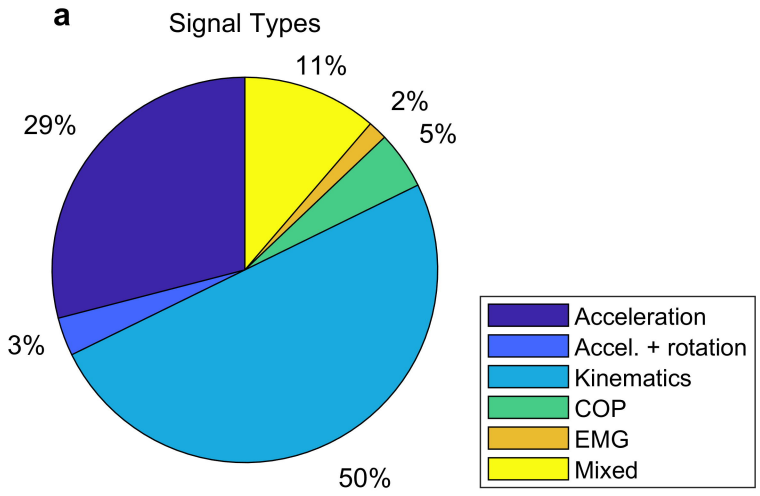

Temporal distribution
a
Number of articles
0
2
4
6
8
2000
2004
2008
2012
2016
2020
2024
Publication year
Study classification
b
Diagnostic
Interventional
Comparative, between-subjects
Comparative, within-subjects
Predictive
Method development
Validation
Reliability
Longitudinal
Clinical application
0
1
20
44
7
26
25
9
1
4
42
3
10
1
2
4
Healthy adults
Diabetic patients
Older adults
Chronic pain patients
Patients with knee osteoarthritis
Fall-prone older adults
Total
Sample sizes (n=63)
c
Sample size (log)
10
20
50
100
200

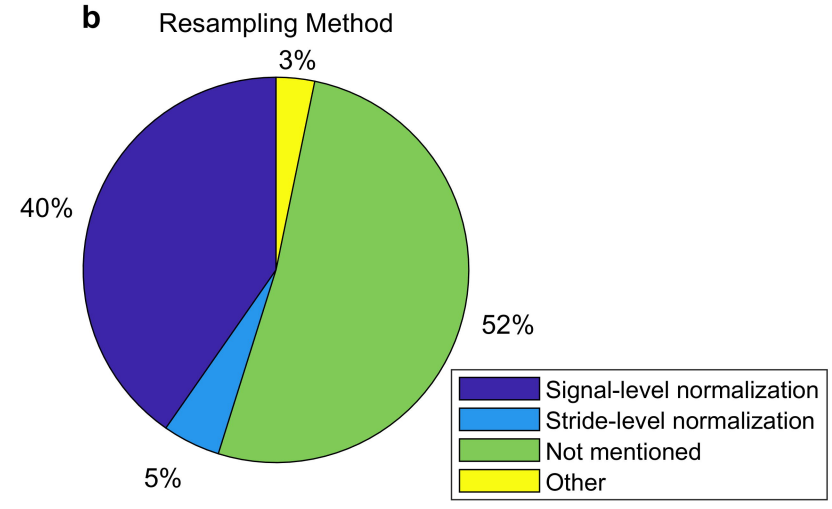
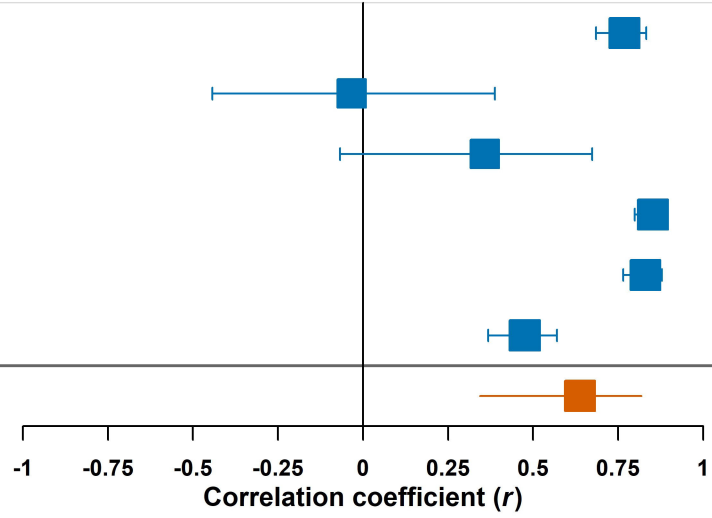

a
Walking Conditions
2%
10%
21%
68%
Overground
Treadmill
Mixed (OG+TW)
Free-living
b
Measurement Systems
15%
5%
50%
31%
3D motion capture
Inertial sensors
Instrumented treadmill
Mixed
c
Anatomical Measurement Sites
5%
5%
18%
55%
18%
Trunk
Lower limbs
Trunk + lower limbs
Foot pressure
Others

Duration of a single analysed walking bout



Category	Time criterion	Stride criterion
Very short	< 2 min	< 100 strides
Short	2 – 5 min	101 – 250 strides
Moderate	5 – 10 min	251 – 500 strides
Long	> 10 min	> 500 strides


a
Signal Types
11%
2%
5%
29%
3%
50%
Acceleration
Accel. + rotation
Kinematics
COP
EMG
Mixed


b
Resampling Method
3%
40%
52%
5%
Signal-level normalization
Stride-level normalization
Not mentioned
Other

Study	N	Subjects	k	Correlation [95% CI]
Jordan 2009	110	11	2	0.77 [0.68, 0.83]
Terrier 2011	20	20	3	-0.03 [-0.44, 0.39]
Terrier 2011	20	20	3	0.36 [-0.07, 0.67]
Terrier 2013	120	20	2	0.85 [0.80, 0.89]
Terrier 2019	108	36	2	0.83 [0.76, 0.88]
Piergiovanni 2024	204	102	2	0.48 [0.37, 0.57]
Overall (CHE + CR2)	**582**	**209**	**14**	**0.64 [0.34, 0.82]**