REVIEW ARTICLE   OPEN

Check for updates

A systematic review on exercise and training-based interventions for freezing of gait in Parkinson's disease

Moran Gilat [1]✉, Pieter Ginis [1], Demi Zoetewei[1], Joni De Vleeschhauwer[1], Femke Hulzinga [1], Nicholas D'Cruz [1] and Alice Nieuwboer[1]

Freezing of gait (FOG) in Parkinson's disease (PD) causes severe patient burden despite pharmacological management. Exercise and training are therefore advocated as important adjunct therapies. In this meta-analysis, we assess the existing evidence for such interventions to reduce FOG, and further examine which type of training helps the restoration of gait function in particular. The primary meta-analysis across 41 studies and 1838 patients revealed a favorable moderate effect size (ES = −0.37) of various training modalities for reducing subjective FOG-severity ($p < 0.00001$), though several interventions were not directly aimed at FOG and some included non-freezers. However, exercise and training also proved beneficial in a secondary analysis on freezers only (ES = −0.32, $p = 0.007$). We further revealed that dedicated training aimed at reducing FOG episodes (ES = −0.24) or ameliorating the underlying correlates of FOG (ES = −0.40) was moderately effective ($p < 0.01$), while generic exercises were not (ES = −0.14, $p = 0.12$). Relevantly, no retention effects were seen after cessation of training (ES = −0.08, $p = 0.36$). This review thereby supports the implementation of targeted training as a treatment for FOG with the need for long-term engagement.

INTRODUCTION

Freezing of gait (FOG) is a very disabling paroxysmal symptom affecting over half of people with Parkinson's disease (PD)[1]. During FOG, patients experience a sudden episodic inability to take an effective step while walking, turning, or initiating gait, leading to a marked reduction or complete cessation in forward progression of the feet despite the intention to walk[1]. FOG episodes are characterized by trembling of the knees, short shuffling steps or complete akinesia, and usually last 1–2 s, although longer periods can occur[1,2]. FOG drastically increases the risk and fear of falling[3].

The most frequent trigger for FOG is turning[4–6], while other common drivers include performing cognitive challenges while walking (i.e., dual-tasking)[5,7], overcoming environmental challenges, such as negotiating doorways[8,9], approaching destinations[6], and reduced visual input, such as when walking in the dark[10]. Greater anxiety has also been related to worse FOG[11].

Like most other symptoms of PD, FOG is more pronounced with advanced disease and when OFF dopaminergic medication[12]. A more specific explanation related to FOG is that episodic shortages of dorsal striatal dopamine in PD will lead to transient epochs of over-activity in the striatal output nuclei (i.e., globus pallidus internus and substantia nigra pars reticulata) and bursts of GABAergic inhibitory projections to the motor thalamus and brainstem locomotor regions inducing gait breakdown[13,14]. Dysfunctional cortical and cerebellar projections to the basal ganglia and brainstem locomotor region may exacerbate the neural manifestation of FOG[15,16].

Figure 1 shows a simplified representation of the brain circuits involved in maintaining gait in healthy adults and PD. As stated above, the pathophysiology of FOG likely involves both localized primary dysfunction of the dorsal motor circuit and a breakdown across compensatory networks[13,14,17–19].

Figure 1a displays the dorsal cortico-basal ganglia and cerebellar motor circuit implicated in the acquisition and execution of automated movements, which in healthy individuals frees up attentional and sensory resources that can then be used to process secondary task demands during walking in usual conditions[20–22]. Specifically, although all these networks interact to some degree when processing varying gait demands, cognitive and sensory input is less essential to maintain a steady gait in healthy subjects.

Figure 1b illustrates that as a result of impaired processing in the dorsal striatum, PD patients become increasingly reliant on external sensory input and cognitive control for their gait, switching to goal-directed behavior[23]. An increased reliance on goal-directed gait control is also evident in healthy older adults[24], but is more pronounced in pwPD, especially in those who experience FOG[25]. Relatively early in the disease, processing across the affected striatum and compensatory networks is capable of controlling gait reasonably well[13,17], although changes in gait metrics can already be detected, such as slower gait speed and increased stride-time variability[25]. At this stage, FOG episodes may occur occasionally and particularly during complex situations whereby the demand for gait control exceeds the combined processing capacity of the motor-cognitive circuits, such as during turning[4]. Aberrant inputs from the limbic circuitry, in particular during high threat conditions[11], may put further strain on striatal processing and thereby exacerbate FOG[13,16,26]. Dopamine replacement therapy and deep brain stimulation (DBS) of the subthalamic nucleus (STN) may at this point still facilitate processing across the dorsal motor circuit, allowing partially automated gait control and thereby reducing the risk for FOG[27,28].

Unfortunately, over time, the inevitable progression of nigral-striatal degeneration will heavily affect processing across the dorsal motor circuit, as shown in Fig. 1c, likely increasing the dependency on compensatory modulation[17,18]. As a result, the higher competition for neural recourses will amplify the risk of transient interference, and as such FOG will emerge more regularly[13,14,17]. Medications and DBS may reduce the frequency

[1]KU Leuven, Department of Rehabilitation Sciences, Neurorehabilitation Research Group (eNRGy), Leuven, Belgium. ✉email: moran.gilat@kuleuven.be
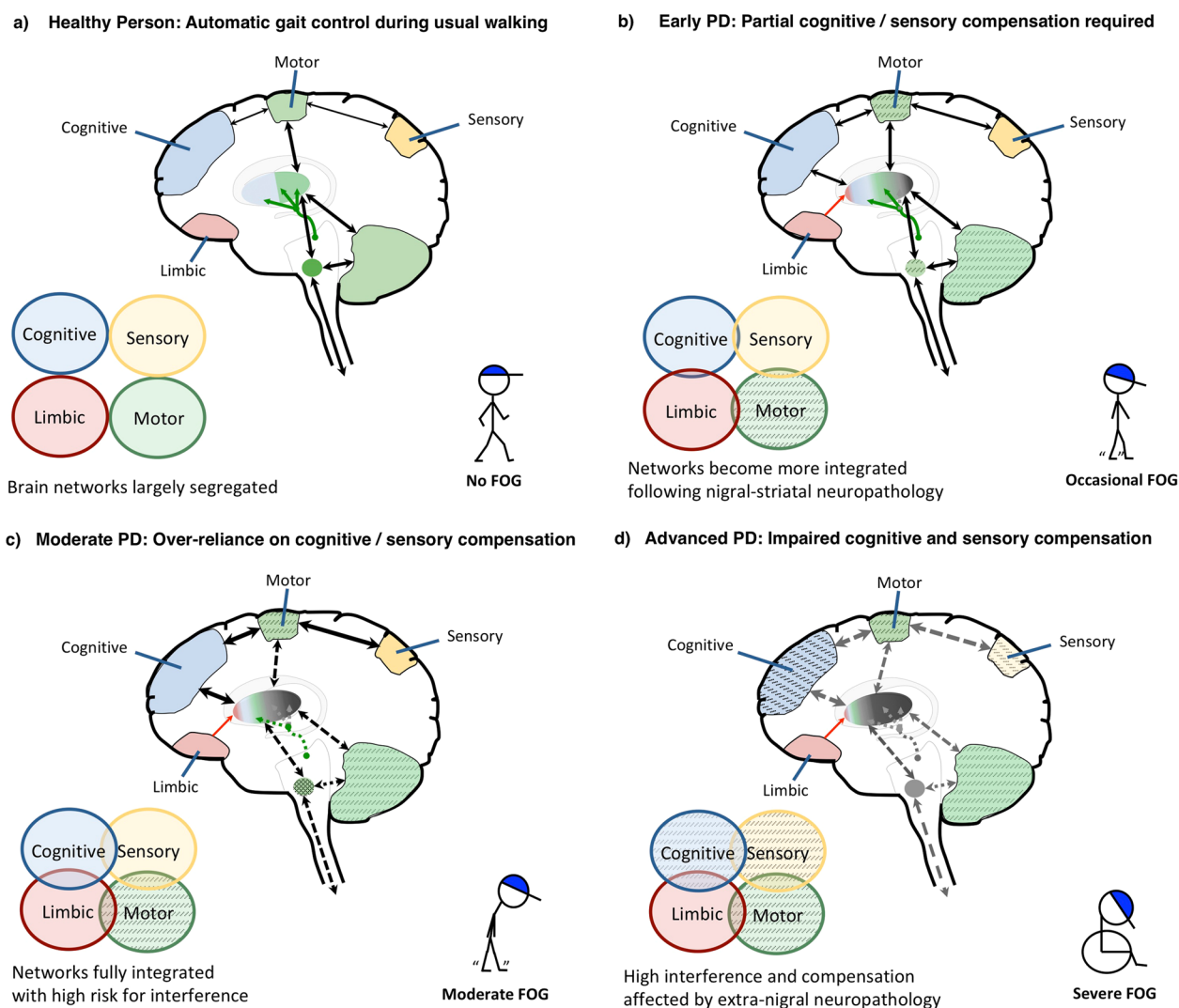
Fig. 1 Theoretical model of gait control in healthy persons and PD patients with FOG. a In healthy individuals, gait automaticity is achieved via processing across dorsal cortico-striatal-thalamo-cerebellar-brainstem neural circuits. Segregation of the motor circuit from other, i.e., sensory, limbic, and cognitive circuits, allows consecutive processing of multiple inputs without interference, ensuring normal gait automaticity. Note, however, that gait is not always automatically controlled. Attentional control will be called upon in gait during complex circumstances. The need for attentional gait control increases with older age and more so in pwPD. **b** Following substantial degeneration of nigral-striatal dopaminergic neurons (indicated by gray coloring in the dorsal striatum) already prominent in early PD, gait automaticity becomes impaired. External sensory input and cognitive control come on to maintain gait control (as indicated by the increase in black arrows, modeling greater inter-circuit connectivity). Occasional FOG occurs when processing demands exceed the combined capacity of motor and compensatory circuits. Limbic input to the striatum (indicated by red arrow) may increase interference and exacerbate FOG. **c** The progression of nigral-striatal neurodegeneration inherent to moderate PD heavily affects processing across the motor circuit (indicated by dotted arrows), increasing the dependency on compensatory circuits. The risk for interference during gait becomes higher, resulting in regular episodes of FOG. **d** In the advanced stages, extra-nigral neuropathology starts to affect processing across the compensatory circuits (indicated by the gray dotted arrows), resulting in severe gait disability and frequent FOG.

of FOG, but are at this point no longer fully adequate for clinical management[19,29].

In the more advanced stages of PD, as shown by Fig. 1d, extra-nigral neuropathology is thought to influence the participation of compensatory gait circuits, which together with the heavily affected motor circuit results in severe gait disability and regular FOG[18]. At this stage, when compensatory gait control starts to fail, dopamine replacement therapies and DBS no longer suffice to treat FOG[29,30].

Clearly, FOG is a highly complex symptom with both motor and non-motor correlates that may be amendable with behavioral therapy. Different types of behavioral interventions can be designed that can safely target FOG at one or multiple of three levels, namely: (1) the circumvention of imminent episodes, such

as via the application of cognitive strategies and/or cueing to induce compensatory gait control; (2) reducing the impact of provocative factors, such as by training pwPD on how to deal with the different triggers of FOG, and; (3) boost compensatory gait control and/or reduce gait interference by targeting the underlying correlates of FOG with the aim to increase resilience against the occurrence of FOG (Fig. 1). However, questions remain on how strong the evidence is for training- and exercise-based interventions to reduce FOG and how a more systematic approach can be facilitated.

Four recent systematic reviews focusing on various forms of exercise incorporated FOG only as a secondary outcome[31–34]. Rutz and Benninger conducted a systematic review on physical therapy specifically for FOG, but included other gait disturbances

as well[35]. Moreover, no meta-analysis was performed to support their conclusions[35].

Two additional systematic reviews were published which addressed whether non-pharmacological interventions had an effect on FOG as a primary outcome. The first summarized evidence on varying therapeutics including non-invasive brain stimulation, and as such did not focus specifically on exercise- or training-based trials[36]. Due to the extensive clinical heterogeneity of the intervention types, a meta-analysis was not feasible. The review also included open-label studies, as well as cross-sectional studies without a training component lasting >1 day, questioning whether the outcomes reflected clinically relevant interventions[36].

The second systematic review selected only RCT studies that represented physiotherapy interventions with FOG as a primary outcome, irrespective of whether it was reported as such in the selected studies[37]. Other strengths were that the authors compared the effects of physical therapy separately for studies with active (i.e., sham intervention) and those with passive (i.e., no treatment) control groups. The outcomes of this meta-analysis revealed that physical therapy improved subjective FOG as compared to both active ($n = 10$, $Z = 3.90$, $p < 0.001$) and passive control groups ($n = 9$, $Z = 3.42$, $p < 0.001$). Long-term retention effects remained significant across eight studies comparing physical therapy to an active control intervention ($Z = 3.89$, $p < 0.001$), in particular following action observation ($n = 4$, $p = 0.002$), but not cueing ($n = 2$, $p = 0.78$). However, Cosentino et al. did not include all types of exercise interventions (e.g., dance and tai-chi), resulting in a total of 19 included studies. Hence, subgroup analyses were low in power and should be interpreted with caution. Finally, the effects of physiotherapy-based interventions that only enrolled PD patients who experience FOG, i.e., freezers, were not examined[37].

Given the complexity of FOG and the involvement of various compensatory circuits and primary motor networks, the present systematic review aims to assess the effects of a broad spectrum of exercise- and training-based interventions on FOG. As exercise compliance and effort is not self-evident in PD[38], this review aims to underpin evidence-based choices to engage in exercise according to patients' preferences. As per Cosentino et al., we will also differentiate between studies with contrasts to either active (i.e., sham) or passive (i.e., no treatment) control groups, and determine the long-term retention effects[37]. Most importantly, to address the specificity of the evidence for FOG, we will investigate studies that enrolled freezers only, given that the more severe disease profiles of freezers could affect effect sizes.

As a secondary aim, we will assess which intervention type may benefit FOG most based on a novel conceptual framework. For this purpose, interventions are split into three subgroups based on their relevance to FOG (Box 1), i.e., A. FOG-specific, B. FOG-relevant, and C. generic exercise. The latter category was included to test the hypothesis that even exercises that are conventionally offered to the wider population for their general health benefits could also benefit FOG in pwPD, as sometimes postulated by the authors of these studies. Critically, the generic exercise interventions did not specifically target FOG as the primary outcome.

Taken together, this systematic review with meta-analysis intends to take a major step forward in determining the evidence for reducing the severity of FOG with various training modes. Based on a critical appraisal of the existing literature, we will provide a comprehensive overview of the benefit of rehabilitation to counter this highly debilitating symptom of PD. As such, we will contribute to a framework for methodical clinical reasoning on how to implement training-interventions over the course of the disease.

SYSTEMATIC SEARCH

Preregistration
The protocol for this systematic review with meta-analysis was prospectively registered and published online on PROSPERO (CRD:42019123882) and can be found in Supplementary Note 1.

Search strategy
Following recent recommendations[39], we searched for literature in the following combination of databases: PubMed, EMBASE, MEDLINE Ovid, Web of Science core collection, and Google Scholar from conception until the 3rd of August 2020. The following combination of search terms was used: Parkinson disease; AND (Freezing OR festination, shuffling); AND (gait OR walking); AND (rehabilitation OR training, exercise, physiotherapy, physical therapy, non-pharmacological, behavioral, cueing, cues, feedback, action observation, dual-task, virtual reality, exergaming, cognitive, auditory, visual, executive, sensory, proprioceptive, imagery, treadmill, wearables, balance, dance, tango, tai chi, strength). The full search criteria are provided in Supplementary Note 1.

Inclusion/exclusion criteria
Literature was selected based on the following a priori inclusion criteria: (i) prospectively collected empirical evidence of any kind on the effect of an exercise or training-based intervention (including cognitive training) of at least 2 days (in order to exclude pre/post studies of a single training session) with a randomized-controlled design; (ii) FOG severity being assessed as an outcome measure; (iii) written in any language and without date restrictions; (iv) article published in a peer-reviewed scientific journal; and (v) evidence based on human participants with a clinical diagnosis of idiopathic PD.

First, non-duplicate titles were screened according to two exclusion criteria: (i) title clearly indicates that the study is a review of the literature with- or without meta-analysis; (ii) title clearly indicates the study is not based on human participants (e.g., animal- or in-vitro studies). The remaining abstracts and full-texts were subsequently screened according to the following exclusion criteria: (i) review of the literature with- or without meta-analysis; (ii) no empirical data; (iii) not based on human participants; (iv) no peer-review; (v) dissertation, conference abstracts or posters; (vi) no participants with idiopathic PD; (vii) no exercise- or training-based interventions (including pharmacological, surgical or other non-behavioral interventions such as non-invasive brain stimulation); (viii) less than 10 participants in total across groups; (ix) no outcome on FOG collected; (x) not randomized-controlled; (xi) FOG only assessed at baseline; (xii) only freezing during movements other than gait (e.g., freezing of the upper limbs, foot-tapping, stationary stepping).

FOG outcomes considered included measures directly related to FOG, such as freezing documented on video, freezing ratios measured with wearable sensors, freezing rated by a clinician, and self-reported FOG using diaries and FOG-related questionnaires. Proxy measures that were indirectly related to FOG, such as measures of gait, balance, falls, activities of daily living, and quality of life were not considered in the present review, because the actual relationship between FOG-severity and these outcomes is indirect and may be unclear.

Literature selection
Two researchers (MG, PG) independently screened non-duplicate titles, abstracts, and full-texts according to the eligibility criteria described above (Fig. 2). In the event of a discrepancy between decisions ($n = 2$), a third moderator (ND) decided on the correct allocation. The same independent procedure was carried out to
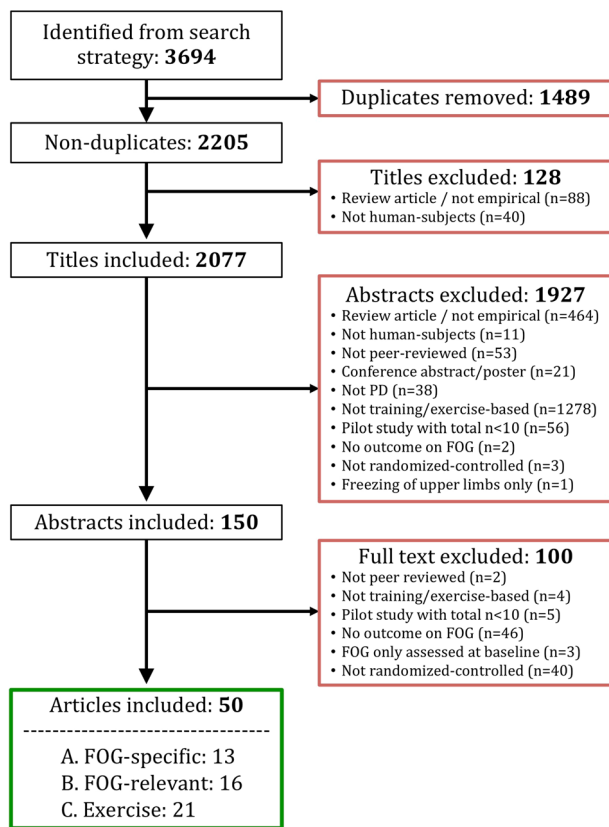
Fig. 2 Flowchart of the systematic article selection strategy. From 3694 articles identified from the search strategy, a total of 50 articles were included for review. Red boxes indicate exclusions of articles during each of the following screening stages: duplicate removal, title, abstract, full-text. Reasons for exclusion are shown inside the red boxes. Green box indicates inclusion.

allocate studies to the conceptual categories displayed in Box 1, whereby discrepancies were resolved by another moderator (AN).

Quality assessment

A modified version of The National Heart, Lung, and Blood Institute Quality Assessment Tool for Controlled Intervention Studies was used to assess the risk of bias within included studies. This scale evaluates the internal validity of controlled studies using 14 predefined criteria. Two assessors (MG and PG) independently scored the criteria.

SYSTEMATIC REVIEW
Literature search results

A total of 2205 non-duplicate titles were screened, resulting in a total of 50 articles eligible for inclusion in the review (Fig. 2). A summary overview of included articles is presented in Table 1 and the full systematic overview of the included articles is presented in Supplementary Table 1.

Study designs

Four studies applied a crossover design[40–43], and the rest a parallel-group design. For crossover trials, only the effects of the first intervention period (i.e., prior to crossover) were considered to minimize potentially confounding carry-over effects.

Box 1 Three conceptual categories of training interventions as based on their relevance to FOG

1. FOG-specific
   Exercise or training-based interventions aimed directly at alleviating imminent FOG episodes or better prepare patients for upcoming FOG while the interventions are applied, and possibly in times beyond. This includes mixed intervention studies, of which at least one training component is directly aimed at reducing FOG episodes or circumventing FOG-provoking circumstances. Examples are: cueing offered to help patients overcome FOG episodes; action-observation training strategies designed to relieve FOG in FOG-provoking situations; and fall-prevention training including strategies to overcome imminent FOG episodes, such as through the use of cueing. FOG was assessed as a primary or other outcome in these studies.

2. FOG-relevant
   Exercise or training-based interventions aimed at training the motor- and/or non-motor correlates of FOG with the aim to reduce the severity or amount of FOG following the intervention, but not aimed at the immediate alleviation of imminent FOG episodes or circumventing FOG-provoking circumstances while the intervention was applied. Examples are cognitive training; cognitive-motor dual-task training; balance training; curved treadmill training; regular treadmill training with cueing that was aimed at improving gait parameters other than FOG; and obstacle avoidance training. FOG was assessed as a primary or other outcome in these studies.

3. Generic exercises
   Conventional physical therapy or generic exercise interventions that are also frequently offered to healthy older individuals to improve physical- and/or mental fitness and other benefits, irrespective of their possible potential to also benefit FOG. Examples are different types of dancing; yoga; physiotherapy not aimed at FOG; aquatic training; Tai-chi; gait training; muscle-power training, and; music therapy. FOG was not assessed as a primary outcome in these studies.

Study Outcome

Eleven studies included FOG as a primary outcome[41,43–52]. Most of these studies were published in recent years (2017 and beyond), with the exception of Pelosin et al.[46] and Fietzek et al.[41]. In the remaining studies, 25 specified FOG as a secondary or tertiary outcome, whilst 14 included FOG as one of the many outcomes without specifying which was considered primary[40,53–63].

The large majority of studies assessed self-reported FOG severity using the "original" Freezing of Gait Questionnaire (FOG-Q; $n = 37$)[64], or the New FOG-Q (NFOG-Q; $n = 10$)[65]. In 2016, the Movement Disorders Task Force deemed the FOG-Q as a recommended instrument for capturing self-perceived FOG[66]. Unlike the NFOG-Q, however, it combines the ratings of gait and FOG. Besides obtaining the FOG-Q, one study also assessed FOG severity using a diary[46]. Two other studies captured the single FOG item of either the Unified Parkinson's Disease Rating Scale Part 2 (UPDRS-II)[53] or the single FOG item of the UPDRS Part 3 (UPDRS-III)[41] besides the FOG-Q. Two studies, however, assessed FOG solely using the single FOG item of the UPRDS-II[67] or UPDRS-III[68]. Four studies also included a semi-objective outcome[69] based on observed FOG severity ratings during the performance of a FOG-provoking walking course[41,54,55,70]. Only Fietzek et al.[41] and Schlenstedt et al.[55] blinded the assessors to group allocation. Two recent studies[43,48] captured objective FOG severity using a FOG-ratio as derived from inertial measurement units during the performance of a turning on the spot task[71]. To date, only one study assessed FOG severity as the percentage of time spent with FOG as measured objectively from video recordings of standardized walking tasks and rated by independent and blinded assessors[50].

Study quality

The full quality assessment is presented in Table 2. All 50 studies randomly allocated participants to one of the study arms (Table 2, item 1). Kalyani et al., however, only randomized the first 75%, and manually allocated the remaining 25% of subjects based on their preference. Their trial should thus be considered

Table 1. Summary overview of included studies.

Study (author, year)	Intervention 1	Intervention 2	N (intervention 1, 2)	%FOG (intervention 1, 2)
Category A: FOG-specific				
Ashburn (2019)[86] & Chivers Seymour (2019)[82 a]	Home-based fall prevention program	Usual care	238, 236	64, 59
Mezzarobba (2017)[44 a]	AO + sonification	Sham without AO	12, 10	100, 100
Pelosin (2018)[45]	AO	Sham without AO	32, 32	100, 100
Agosta (2017)[53]	AO	Sham without AO	12, 13	100, 100
Cui (2017)[87]	PT with cueing	Sham PT without cueing	20, 20	NR, NR
Ginis (2016)[70]	Corrective feedback on gait performance	Sham gait training without corrective feedback	20, 18	70, 55.6
Canning (2015)[74]	Fall prevention exercises	Usual care	115, 116	46, 53
Martin (2015)[40]	Cueing training	Wait-list control	10, 9	100, 100
Fietzek (2014)[41]	Cueing training	Wait-list control	14, 8	100, 100
Allen (2010)[101]	Fall prevention exercises	Usual care	24, 24	NR, NR
Pelosin (2010)[46]	AO	Sham without AO	9, 9	100, 100
Nieuwboer (2007)[42 a]	Cueing training at home	Wait-list control	76, 77	20.3, 20.9
Category B: FOG-relevant				
Bekkers (2020)[47 a]	TT with virtual reality	Sham TT without virtual reality	34, 43	100, 100
King (2020)[43]	Agility bootcamp-cognitive program	Education and relaxation control group	23, 19	100, 100
Silva-Batista (2020)[48]	Motor-cognitive balance training	Sham traditional motor rehabilitation	17, 15	100, 100
Capecci (2019)[114]	Robot-assisted TT	Sham TT without robot	48, 48	68, 63
Clerici (2019)[49]	Motor-cognitive training + aquatic training	Sham motor-cognitive training only	27, 25	100, 100
Wróblewska (2019)[54]	Nordic walking training	Usual care	20, 20	100, 100
Schlenstedt (2018)[55]	Resistance training	Balance training	12, 8	100, 100
Walton (2018)[50]	Cognitive training specific to FOG	Sham cognitive training not specific to FOG	20, 18	100, 100
Zhu (2018)[51]	Aquatic obstacle training	Sham aquatic training without obstacles	23, 23	NR, NR
Cheng (2017)[52]	Curved walking TT	Sham trunk and upper limb exercises	12, 12	NR, NR
Santos (2017b)[100]	Slack-line training	NR	11, 11	NR, NR
Schlick (2016)[56]	TT with visual cueing not targeted at FOG	Sham TT without visual cueing	10, 10	NR, NR
King (2015)[79]	Individually-supervised (1) or group-supervised (2) agility boot camp training	Unsupervised home-based agility boot camp training (3)	21, 20, 17	46, 46, 60
Ricciardi (2015)[80 b]	PT for most affected (1), or least affected body side (2)	PT for both body sides (3)	9, 9, 10	NR, NR, NR
Kadivar (2011)[83 a]	Auditory cued stepping in place	Sham internally cued stepping in place	8, 8	50, 37.5
Frazzitta (2009)[57]	TT and cueing training	Sham cueing training without TT	20, 20	100, 100
Category C: Generic exercise				
Kalyani (2020)[58]	Dancing seated and standing	Usual care	17, 16	NR, NR
Pohl (2020)[90]	Group- and music-based Ronnie Gardiner method	Usual care	26, 20	NR, NR
Hubble (2019)[84]	Trunk exercises + fall education brochures	Sham fall education	11, 11	NR, NR
Medijainen (2019)[115]	PT including cueing not specific to FOG	Wait-list control	12, 12	NR, NR
Rocha (2018)[73]	Tango	Mixed genre dancing	10, 11	NR, NR
Sedaghati (2018)[59]	Alexander-based corrective techniques on forward flexed posture	NR	13, 13	NR, NR
Van Puymbroeck (2018)[60]	Yoga	Wait-list control	15, 12	NR, NR
Carpinella (2017)[75]	PT with biofeedback (Gamepad system)	Sham PT without biofeedback	17, 20	NR, NR
Carroll (2017)[81]	Aquatic training	Usual care	10, 8	NR, NR
Santos (2017)[61]	Progressive resistance training	Usual care	13, 15	NR, NR

M. Gilat et al.

Table 1 continued

Study (author, year)	Intervention 1	Intervention 2	N (intervention 1, 2)	%FOG (intervention 1, 2)
Xiao (2017)[78]	Tai-Chi ball exercises	Usual care	25, 25	NR, NR
Byl (2015)[85] a	Gait training with visual and kinesthetic feedback	Sham gait training without feedback	7, 5	NR, NR
Romenets (2015)[98]	Tango	Waitlist control + pamphlet for PD exercises at home	18, 15	NR, NR
Duncan (2014)[76] c	Tango	Usual care	5, 5	NR, NR
Paul (2014)[116]	Muscle power training	Sham low-intensity exercises at home	20, 20	NR, NR
Volpe (2013)[72]	Irish set dancing	PT with cueing not specific to FOG	12, 12	NR, NR
Duncan (2012)[77] a	Tango	Usual care	26, 26	NR, NR
Reuter (2011)[68]	Nordic walking (1), or regular walking (2) training	Flexibility and relaxation training (3)	30, 30, 30	NR, NR, NR
Hackney (2009)[62]	Tango (1) or Waltz/Foxtrot (2)	Usual care (3)	14, 17, 17	57.1, 52.9, 29.4
Hackney (2007)[63]	Tango	Sham strength and flexibility exercises	9, 10	NR, NR
Pacchetti (2000)[67]	Music therapy (i.e., use of musical instruments)	PT including gait and balance training	16, 16	NR, NR

Studies are grouped per intervention category as highlighted in bold font.

AO Action observation, FOG freezing of gait, PT physical therapy, TT Treadmill training, N number of subjects analyzed in each group for the FOG outcome at the primary endpoint; %FOG Percentage of subjects that were classified as being "freezers" (i.e., people with Parkinson's disease who experience FOG) at baseline, NR Not reported or values could not be computed with the information provided.

aFOG meta-data provided by authors upon email request.

bAuthors contacted by email to retrieve FOG data for inclusion in the meta-analysis.

cSame participants as Duncan and Earhart (2012)[77].

pseudo-randomized[58]. A total of 36 studies reported adequate concealment of treatment allocation, mostly by using computerized randomization procedures by an independent investigator. Ten studies reported that participants were kept blinded as much as possible by not informing them of the study aims (Table 2, item 4). Thirty-three studies kept the assessors of study outcomes blinded to treatment allocation (Table 2, item 5). Two studies also blinded the treatment providers by involving physiotherapist[72] or dance teachers[73] to provide usual care without being told the aims of the study and group assignment of the participants. The large majority of studies ($n = 36$) reported less than 20% overall dropout. Seven studies reported more than 20% overall dropout across the interventions, whilst another six did not report the dropout rates (Table 2, item 7). The dropout rates were similar between groups in three of the seven studies (Table 2, item 8), whereas the dropout rate was higher in the control group for two studies[40,50] and higher in the intervention group for two other studies[74,75]. Adherence rates were high (70–100%) in all of the 31 studies reporting compliance (Table 2, item 9). Only twenty studies reported to have conducted an a priori power calculation to determine the required sample size, and of these, six did not reach their recruitment target (Table 2, item 12). Adequate statistical power in the analysis can therefore only be assumed for 14 of the 50 included studies. As can be seen in Table 2, the large majority ($n = 43$) of studies had a high or unclear risk of bias on three or more items on this quality assessment scale.

Type of interventions

Table 1 presents the precise intervention types offered in each study. Eight studies trialed a dancing-based intervention, seven a cueing-based intervention, four applied action-observation training, one cognitive training, and the remaining studies investigated different types of physical therapy such as treadmill, aquatic, Nordic walking, tai-chi, balance, resistance, yoga, slackline, curved walking, and fall prevention exercises. For the purpose of this review, the interventions were divided into three newly devised subcategories based on their relevance to FOG (Box 1). A total of 13 studies were categorized as being specifically designed to reduce FOG episodes, 16 as being designed to ameliorate the correlates underlying FOG, and 21 as being non-specific and unrelated to FOG (Table 1).

Dosage of interventions

The duration of the interventions ranged between 2 weeks and 2 years, with a median of 8 weeks (Supplementary Table 1). The dosage (number of weeks * number of sessions per week * time per session) of supervised interventions ranged from 3 to 224 h with a median of 15 h. The total dosage, including unsupervised (i.e., home-based) sessions also ranged from 3 to 224 h, but with a median of 18 h. In studies, which did not report the unsupervised adherence rates, we assumed a 1× per week adherence of similar duration as instructed by the investigators.

Medication status

In 38 studies, which provided training in the ON medication state, the dopaminergic medication status of participants was reported. A total of 44 studies reported the medication status of participants during the assessments. Of these, 39 studies assessed participants during the ON state, four[43,76,77,78] in the practically defined "OFF" state, and one study[50] assessed participants both "ON" and "OFF" dopaminergic medications on two separate testing days (Supplementary Table 1).

Control groups

Twenty-two studies compared the effects of the intervention against a passive control group (i.e., delayed-start or usual care

Table 2. Quality assessment within studies per category.

Study (author, year)	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A. FOG-specific														
Ashburn, 2019					X		X						X	
Chivers Seymour, 2019					X							X		
Mezzarobba, 2018					X			X	X				X	X
Pelosin, 2018					X	X				X		X	X	X
Agosta, 2017					X			X	X			X	X	X
Cui, 2017				X	X	X			X	X		X	X	X
Ginis, 2016				X	X	X						X		X
Canning, 2015					X			X	X		X			
Martin, 2015				X	X	X		X		X	X	X	X	X
Fietzek, 2014					X		X			X		X	X	X
Allen, 2010					X					X			X	
Pelosin, 2010				X								X	X	X
Nieuwboer, 2007					X							X		
B. FOG-relevant														
Bekkers, 2020					X							X	X	X
King, 2020					X		X					X	X	X
Silva-Batista, 2020	X													X
Capecci, 2019					X									X
Clerici, 2019												X		
Wroblewska, 2019					X	X	X	X	X			X	X	X
Schlenstedt, 2018				X	X			X	X	X		X	X	X
Walton, 2018							X	X	X		X			
Zhu, 2018									X	X		X	X	X
Cheng, 2017					X									X
Santos, 2017b				X	X							X	X	X
Schlick, 2016				X	X	X			X	X		X	X	X
King, 2015					X						X			X
Ricciardi, 2015							X	X	X	X		X	X	X
Kadivar, 2011		X	X	X	X	X				X			X	X
Frazzitta, 2009		X		X	X				X	X		X	X	
C. Generic exercise														
Kalyani, 2020		X	X	X	X	X							X	
Pohl, 2020					X					X		X		X
Hubble, 2019				X	X		X	X		X				X
Medijainen, 2019		X	X	X						X		X	X	X
Rocha, 2018				X			X			X		X		X
Sedaghati, 2018		X	X	X	X	X	X	X	X	X		X	X	X
Van Puymbroeck, 2018					X		X			X	X	X	X	X
Carpinella, 2017					X		X	X	X	X	X	X	X	X
Carroll, 2017					X							X	X	
Santos, 2017a					X	X				X	X	X	X	X
Xiao, 2017		X	X	X	X	X	X	X	X			X	X	X
Byl, 2015		X	X	X	X	X				X		X	X	X
Romenets, 2015				X	X		X			X		X	X	
Duncan, 2014				X	X			/	/	X		X	X	X
Paul, 2014						X				X			X	
Volpe, 2013					X					X		X	X	X
Duncan, 2012				X	X			X			X	X		
Reuter, 2011					X							X	X	X

Table 2 continued

Study (author, year)	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Hackney, 2009												X		X
Hackney, 2007		X		X								X	X	
Pachetti, 2000				X			X	X	X			X	X	X	X

Studies are grouped per intervention category as highlighted in bold font. *Empty cell* low risk of bias, *Cell containing a cross* high or unclear risk of bias, / not Applicable. The following criteria were scored:
(1) Was the study described as randomized, a randomized trial, a randomized clinical trial, or an RCT?
(2) Was the method of randomization adequate (i.e., use of randomly generated assignment)?
(3) Was the treatment allocation concealed (so that assignments could not be predicted)?
(4) Were study participants kept blinded to the expected effects of the intervention?
(5) Were the people assessing the outcomes blinded to participants' group assignments?
(6) Were the groups similar at baseline on important characteristics that could affect outcomes (e.g., demographics, risk factors, comorbid conditions)?
(7) Was the overall dropout rate from the study at endpoint 20% or lower of the number allocated to treatment?
(8) Was the differential dropout rate (between treatment groups) at endpoint 15 percentage points or lower?
(9) Was there high adherence (>75%) to the intervention protocols for each treatment group?
(10) Were other interventions avoided or similar in the groups (e.g., similar background treatments)?
(11) Were outcomes assessed using valid and reliable measures, implemented consistently across all study participants?
(12) Did the authors report that the sample size was sufficiently large to be able to detect a difference in the main outcome between groups with at least 80% power?
(13) Were outcomes reported or subgroups analyzed pre-specified (i.e., identified before analyses were conducted)?
(14) Were all randomized participants analyzed in the group to which they were originally assigned, i.e., did they use an intention-to-treat analysis?

control group without a training intervention designed to induce any measurable training effects on the outcomes of interest). Twenty-five studies compared the intervention group against an active (e.g., sham interventions or standard physiotherapy) control group. Finally, three studies compared the effects between two or more intervention groups and were thus not included in the meta-analysis. Specifically, King et al. compared the effects of a home-based exercise program with that of an individual- as well as a group-based exercise program[79]. The intervention contained exercises aimed at the underlying correlates of FOG (category B). Ricciardi et al. compared the effects across three physiotherapy-based programs, namely one aimed at improving the worst body-side of PD, one aimed at improving the best-side of PD, and one offering standard physiotherapy. The investigators aimed to improve step asymmetry underlying FOG (category B)[80]. Finally, Rocha et al. compared Argentinian tango dancing to mixed-genre dancing (category C)[73].

Study populations
Characteristics of the study populations, involving 2972 patients, are detailed in Supplementary Table 1. The number of participants in each group ranged from 5 to 238, with a median number of 17 subjects in each group. Across all active groups, 61% of participants were male and 39% female and in the control groups 55% were male and 45% female. Three studies did not report the participant's sex[43,68,78]. The mean (SD) age across all groups was

68.8 (3.8) years, across the intervention groups 68.6 (3.7) and control groups 69.0 (4.0), ranging between 60 and 81 years on group average. A total of 45 studies reported disease duration as the number of years since clinical diagnosis. The mean (SD) disease duration across all groups was 8.4 (2.3) years, across the intervention groups 8.1 (2.2) and control groups 8.7 (2.4), ranging between 3 and 13 years on group average. A total of 42 studies reported the Hoehn and Yahr stages (HY). The median HY stage across all groups was 2.5, ranging between 1.6 and 3.2 on group average. A total of 40 studies reported the scores on the motor part of the Movement Disorders Society Unified Parkinson's Disease Rating scale (UPDRS-III) during the dopaminergic ON state. The mean (SD) UDPRS-III across all groups was 28.3 (8.0), across the intervention groups 28.2 (8.6) and control groups 28.4 (8.2), ranging between 9.2 and 51.6 on group average. Twenty-four studies reported the group averages on the Mini-Mental State Examination (MMSE) as a rapid cognitive screening test. The mean (SD) MMSE across all these groups was 27.8 (1.1), across the intervention groups 27.7 (1.0) and control groups 27.8 (1.1), ranging from 25.5 to 29.3 on group average. Only eight studies reported the Montreal Cognitive Assessment (MoCA) as a rapid screening test that is more sensitive to the cognitive deficits in PD. The mean (SD) MoCA across the groups in these studies was 25.8 (1.0). The mean (SD) MoCA of the intervention groups was 25.9 (1.0) and of the control groups 25.6 (1.1).

META ANALYSIS

Data extraction

Two researchers (MG, PG) independently extracted relevant metadata from the included studies using a standardized score form, see Supplementary Note 1. Inconsistencies in data entries were screened for and resolved by the researchers prior to conducting the analyses. Given that the large majority of studies assessed FOG subjectively using questionnaires, only data obtained from the FOG-Q and NFOG-Q were entered into the meta-analysis. In one study, median (interquartile-range; IQR) values were reported and due to the relatively low sample size of that study ($n = 10$ and $n = 8$ per intervention) mean (SD) values could not be imputed[81]. As such, the median (IQR) values were entered into the analysis despite the likelihood of skewness in the study data. A sensitivity analysis was performed without this particular study, showing that the results of the main analysis barely changed (primary analysis: $\Delta Z = -0.06$, $\Delta I^2 = 2\%$), and as such, the data of this study were entered into the analyses. Ginis et al.[70] and Bekkers et al.[47] reported the NFOG-Q for a subgroup of PD patients who experienced FOG in their study. In the event of missing data on FOG ($n = 9$), the corresponding authors were contacted via email with the request to enter the missing items in a standardized data collection form provided. Six of the nine contacted authors provided the missing data and could be included in the meta-analysis[43,44,47,77,82,83], while three did not reply[80,84,85]. The data of Chivers Seymour et al.[82] are the same as Ashburn et al.[86] and the data of Duncan and Earhart[77] are the same as Duncan and Earhart[76]. Hence these data were only entered once. From the Chivers Seymour et al. data set[82], only the outcomes obtained from the PD subgroup that were classified as being freezers at baseline and for whom NFOG-Q data was available at both baseline and the primary endpoint (6 months) were included in the meta-analyses (intervention $n = 80$, control $n = 79$). Data from a subsample of included studies ($n = 15$) that also completed the (N)FOG-Q at follow-up were entered into the secondary retention analysis (intervention $n = 257$, control $n = 263$), As described in the Secondary analyses outcomes section below. Walton et al.[50], Reuter et al.[68] and Pacchetti et al.[67] did not collect the (N)FOG-Q and could thus not be entered into the meta-analysis.

Data synthesis

Data from the FOG-Q and NFOG-Q were standardized and pooled using an inverse-variance random-effects meta-analysis in RevMan (v5.3). First, mean (N)FOG-Q values at baseline were compared between intervention and control groups for each study using independent sample t-tests. This analysis revealed that the intervention and control groups within each study had a comparable mean (N)FOG-Q scores at baseline (all $p > 0.05$), except for the study of Wroblewska et al.[54]. As such, an analysis of final measures was employed for the meta-analysis, whereby the standardized means of the post-intervention (N)FOG-Q scores were calculated and entered into the meta-analysis with 95% confidence intervals and two-sided p-values. In the study of Wroblewska et al.[54], the intervention group receiving Nordic walking training started with a higher mean (SD) FOG-Q score of 13.8 (2.3) than the passive no-intervention control group (9.3 (1.8), $p < 0.01$). Therefore, we also performed sensitivity analyses after the removal of this particular study. In another study, two intervention groups were compared against a control group[62]. As such, the mean FOG-Q scores of both intervention groups were entered into the Revman calculator and their combined standardized mean score was compared against the control group. For all analyses, heterogeneity between the studies in effect measures was explored using the χ^2 and I^2 statistics, whereby a significant χ^2 statistic ($p < 0.05$) and/or an I^2 value >50% was considered representative of substantial heterogeneity. In the event of substantial heterogeneity, a leave-one-out sensitivity analysis was conducted. Funnel plots were computed to evaluate potential bias. Studies were considered outliers if their effect estimates fell outside the 95% confidence interval of the pooled effect estimates, visualized as dotted lines in the Funnel plots. These are presented in Supplementary Data 1.

Primary analysis outcome

The primary objective was to assess the effects of any exercise or training-based intervention on FOG in PD as compared to a control intervention. The primary meta-analysis was therefore conducted across all studies reporting either the FOG-Q or NFOG-Q as an outcome ($n = 41$ studies, totaling $N = 933$ and $N = 905$ subjects in the intervention and control groups, respectively). The test for overall effect revealed a favorable moderate effect for exercise/training compared to any type of control group ($Z = 4.91$, $p < 0.00001$), effect size (ES, [95% confidence intervals] = -0.37 [-0.51, -0.22]), but with large statistical heterogeneity across study effects ($\chi^2(40) = 84.19$, $p < 0.01$ and $I^2 = 52\%$). The forest plot of the primary analysis is presented in Fig. 3. The sensitivity analysis revealed that the heterogeneity assessed with the χ^2 statistic became non-significant ($p > 0.05$) after removal of the outlier studies: Wroblewska et al.[54], Volpe et al.[72], and Cui et al.[87]. After excluding these three studies, 38 studies ($N_{active} = 881$, $N_{control} = 853$) remained, which revealed that the overall effect was still significant ($Z = 4.60$, $p < 0.00001$, ES = -0.24 [-0.35, -0.14]), but now with little statistical heterogeneity across study effects ($\chi^2(37) = 40.48$, $p = 0.32$ and $I^2 = 9\%$). Funnel plots are presented in Supplementary Data 1.

Available data of two studies could not be entered into the meta-analysis. Walton et al. conducted a cognitive training intervention designed to target the underlying correlates of FOG (i.e., attention and executive functioning) and compared its effects to that of a control cognitive training aimed at memory functioning, which is considered unrelated to FOG. Importantly, this was the only RCT to capture the percentage of time spent with FOG (%FOG) from video footage of a standardized gait assessment as the primary outcome[50,88]. Cognitive training targeting FOG reduced FOG severity during the ON medication

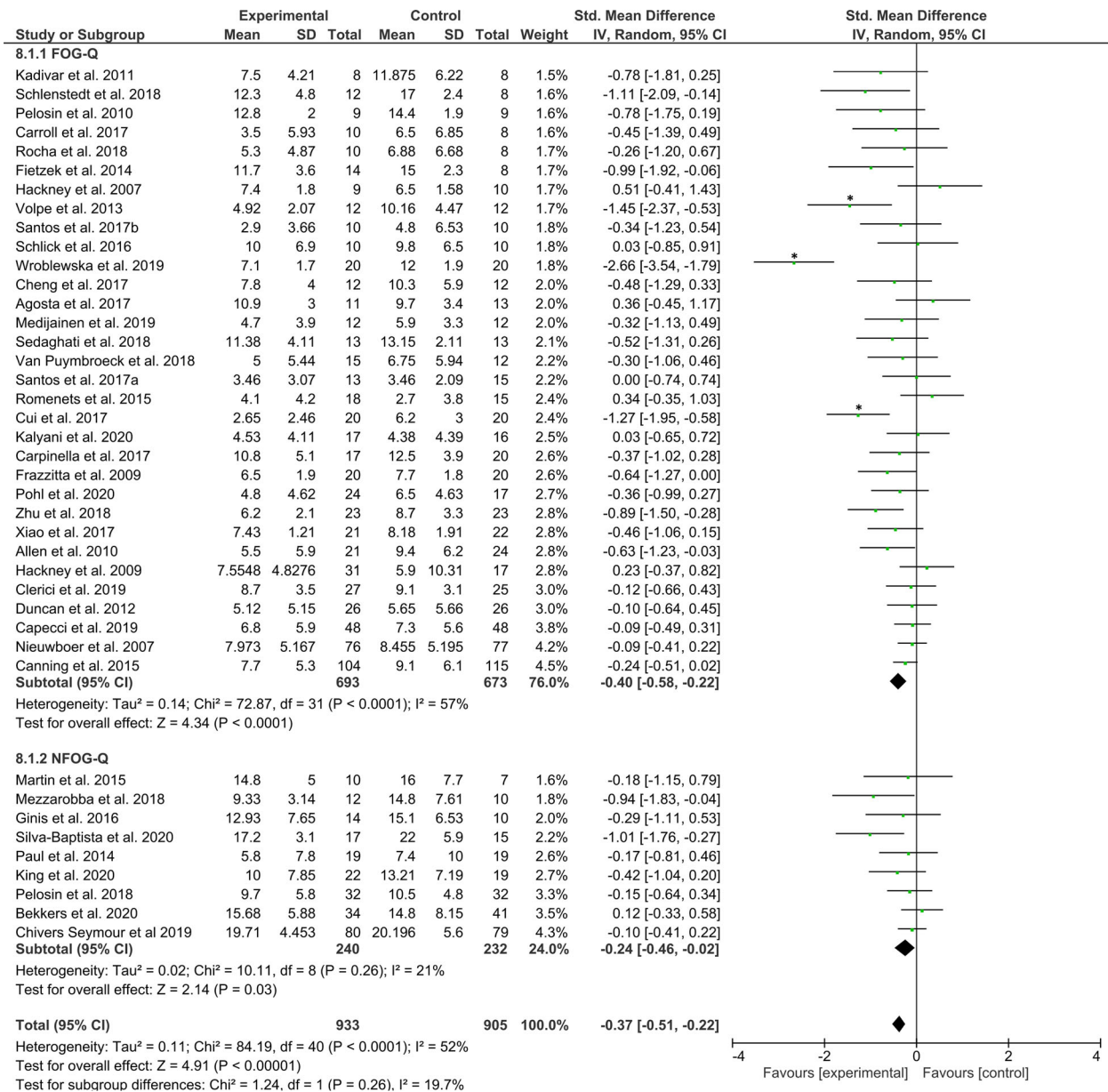| Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
8.1.1 FOG-Q								
Kadivar et al. 2011	7.5	4.21	8	11.875	6.22	8	1.5%	-0.78 [-1.81, 0.25]
Schlenstedt et al. 2018	12.3	4.8	12	17	2.4	8	1.6%	-1.11 [-2.09, -0.14]
Pelosin et al. 2010	12.8	2	9	14.4	1.9	9	1.6%	-0.78 [-1.75, 0.19]
Carroll et al. 2017	3.5	5.93	10	6.5	6.85	8	1.6%	-0.45 [-1.39, 0.49]
Rocha et al. 2018	5.3	4.87	10	6.88	6.68	8	1.7%	-0.26 [-1.20, 0.67]
Fietzek et al. 2014	11.7	3.6	14	15	2.3	8	1.7%	-0.99 [-1.92, -0.06]
Hackney et al. 2007	7.4	1.8	9	6.5	1.58	10	1.7%	0.51 [-0.41, 1.43]
Volpe et al. 2013	4.92	2.07	12	10.16	4.47	12	1.7%	-1.45 [-2.37, -0.53]
Santos et al. 2017b	2.9	3.66	10	4.8	6.53	10	1.8%	-0.34 [-1.23, 0.54]
Schlick et al. 2016	10	6.9	10	9.8	6.5	10	1.8%	0.03 [-0.85, 0.91]
Wroblewska et al. 2019	7.1	1.7	20	12	1.9	20	1.8%	-2.66 [-3.54, -1.79]
Cheng et al. 2017	7.8	4	12	10.3	5.9	12	2.0%	-0.48 [-1.29, 0.33]
Agosta et al. 2017	10.9	3	11	9.7	3.4	13	2.0%	0.36 [-0.45, 1.17]
Medijainen et al. 2019	4.7	3.9	12	5.9	3.3	12	2.0%	-0.32 [-1.13, 0.49]
Sedaghati et al. 2018	11.38	4.11	13	13.15	2.11	13	2.1%	-0.52 [-1.31, 0.26]
Van Puymbroeck et al. 2018	5	5.44	15	6.75	5.94	12	2.2%	-0.30 [-1.06, 0.46]
Santos et al. 2017a	3.46	3.07	13	3.46	2.09	15	2.2%	0.00 [-0.74, 0.74]
Romenets et al. 2015	4.1	4.2	18	2.7	3.8	15	2.4%	0.34 [-0.35, 1.03]
Cui et al. 2017	2.65	2.46	20	6.2	3	20	2.4%	-1.27 [-1.95, -0.58]
Kalyani et al. 2020	4.53	4.11	17	4.38	4.39	16	2.5%	0.03 [-0.65, 0.72]
Carpinella et al. 2017	10.8	5.1	17	12.5	3.9	20	2.6%	-0.37 [-1.02, 0.28]
Frazzitta et al. 2009	6.5	1.9	20	7.7	1.8	20	2.6%	-0.64 [-1.27, 0.00]
Pohl et al. 2020	4.8	4.62	24	6.5	4.63	17	2.7%	-0.36 [-0.99, 0.27]
Zhu et al. 2018	6.2	2.1	23	8.7	3.3	23	2.8%	-0.89 [-1.50, -0.28]
Xiao et al. 2017	7.43	1.21	21	8.18	1.91	22	2.8%	-0.46 [-1.06, 0.15]
Allen et al. 2010	5.5	5.9	21	9.4	6.2	24	2.8%	-0.63 [-1.23, -0.03]
Hackney et al. 2009	7.5548	4.8276	31	5.9	10.31	17	2.8%	0.23 [-0.37, 0.82]
Clerici et al. 2019	8.7	3.5	27	9.1	3.1	25	3.0%	-0.12 [-0.66, 0.43]
Duncan et al. 2012	5.12	5.15	26	5.65	5.66	26	3.0%	-0.10 [-0.64, 0.45]
Capecci et al. 2019	6.8	5.9	48	7.3	5.6	48	3.8%	-0.09 [-0.49, 0.31]
Nieuwboer et al. 2007	7.973	5.167	76	8.455	5.195	77	4.2%	-0.09 [-0.41, 0.22]
Canning et al. 2015	7.7	5.3	104	9.1	6.1	115	4.5%	-0.24 [-0.51, 0.02]
Subtotal (95% CI)			**693**			**673**	**76.0%**	**-0.40 [-0.58, -0.22]**
Heterogeneity: Tau² = 0.14; Chi² = 72.87, df = 31 (P < 0.0001); I² = 57%								
Test for overall effect: Z = 4.34 (P < 0.0001)								
8.1.2 NFOG-Q								
Martin et al. 2015	14.8	5	10	16	7.7	7	1.6%	-0.18 [-1.15, 0.79]
Mezzarobba et al. 2018	9.33	3.14	12	14.8	7.61	10	1.8%	-0.94 [-1.83, -0.04]
Ginis et al. 2016	12.93	7.65	14	15.1	6.53	10	2.0%	-0.29 [-1.11, 0.53]
Silva-Baptista et al. 2020	17.2	3.1	17	22	5.9	15	2.2%	-1.01 [-1.76, -0.27]
Paul et al. 2014	5.8	7.8	19	7.4	10	19	2.6%	-0.17 [-0.81, 0.46]
King et al. 2020	10	7.85	22	13.21	7.19	19	2.7%	-0.42 [-1.04, 0.20]
Pelosin et al. 2018	9.7	5.8	32	10.5	4.8	32	3.3%	-0.15 [-0.64, 0.34]
Bekkers et al. 2020	15.68	5.88	34	14.8	8.15	41	3.5%	0.12 [-0.33, 0.58]
Chivers Seymour et al 2019	19.71	4.453	80	20.196	5.6	79	4.3%	-0.10 [-0.41, 0.22]
Subtotal (95% CI)			**240**			**232**	**24.0%**	**-0.24 [-0.46, -0.02]**
Heterogeneity: Tau² = 0.02; Chi² = 10.11, df = 8 (P = 0.26); I² = 21%								
Test for overall effect: Z = 2.14 (P = 0.03)								
Total (95% CI)			**933**			**905**	**100.0%**	**-0.37 [-0.51, -0.22]**
Heterogeneity: Tau² = 0.11; Chi² = 84.19, df = 40 (P < 0.0001); I² = 52%								
Test for overall effect: Z = 4.91 (P < 0.00001)								
Test for subgroup differences: Chi² = 1.24, df = 1 (P = 0.26), I² = 19.7%								

Favours [experimental] Favours [control]

Fig. 3 Primary analysis comparing all types of training/exercise interventions against any type of control group. The studies are arranged per weight of the study determined by sample size; *significant outliers removed in the sensitivity analysis.

state in those who experienced FOG during the baseline assessment ($N_{intervention} = 20$, $N_{control} = 18$, $p = 0.002$)[50]. This indicates that cognitive training may help to increase the processing capacity of the compensatory circuits involved in PD gait control, thereby reducing the amount of FOG (Fig. 1). No such improvement was seen during the OFF medication state ($p = 0.80$)[50]. King et al. compared the effects of different delivery modes of the same physical therapy intervention (i.e., Agility Boot Camp exercise program). Delivery was provided via an unsupervised home-based program ($n = 17$) for 4 weeks, or by individual- ($n = 21$) or group-class programs ($n = 20$), supervised 3× per week for 4 weeks. Their investigations revealed that group-based supervised delivery successfully reduced subjective FOG ($p < 0.01$), whereas both individually supervised ($p = 0.431$) and unsupervised home-based delivery ($p = 0.308$) did not. Not all patients had FOG, but the number of freezers in each group was roughly equivalent (46% for individual- and group-classes vs. 60% in the home-based group)[79].

Secondary analyses outcomes

Secondary analyses were conducted to explore the main findings. Firstly, to indicate that training/exercise interventions are superior in reducing FOG severity to no intervention, a meta-analysis was conducted on 21 studies that compared the effects of training/ exercise ($N = 576$) against a passive (i.e., wait-list, delayed-start, usual care) control group ($N = 551$) whom received no intervention. The test for overall effect revealed a significant moderate effect favoring the interventions ($Z = 3.38$, $p = 0.0007$, ES $= -0.36$ [-0.57, -0.15]), but again with high levels of statistical heterogeneity across study effects ($\chi^2(20) = 51.65$, $p < 0.001$, $I^2 = 61\%$). The Forest plot is presented in Fig. 4. A sensitivity analysis revealed that the statistical heterogeneity dropped to ($\chi^2(18) = 13.3$, $p = 0.77$, $I^2 = 0\%$) after excluding the studies by Wroblewska et al.[54] and Cui et al.[87], while the test for overall effects remained significant (19 studies, $N_{intervention} = 536$, $N_{control} = 511$, $Z = 3.18$, $p = 0.001$, ES $= -0.20$ [-0.32, -0.08]). Funnel plots for each of the secondary analyses are presented in Supplementary Data 1.
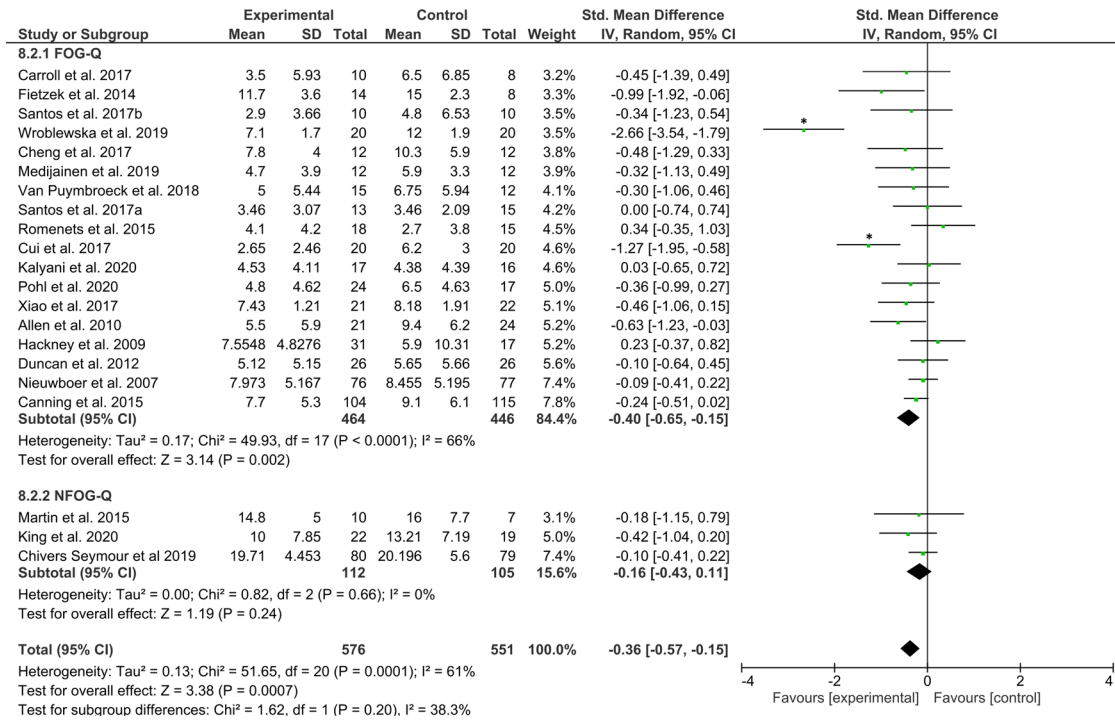
Fig. 4 Secondary analysis comparing the effect of all types of training/exercise interventions against passive control groups. The studies are arranged per weight of the study as based on sample size; *significant outliers removed in the sensitivity analysis.
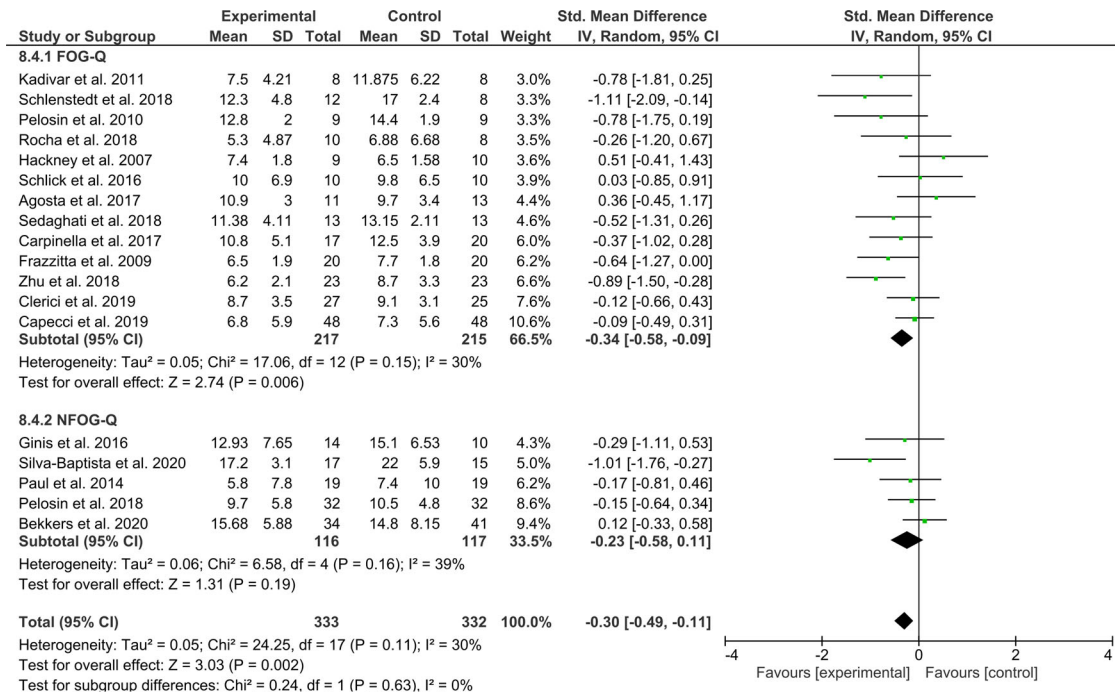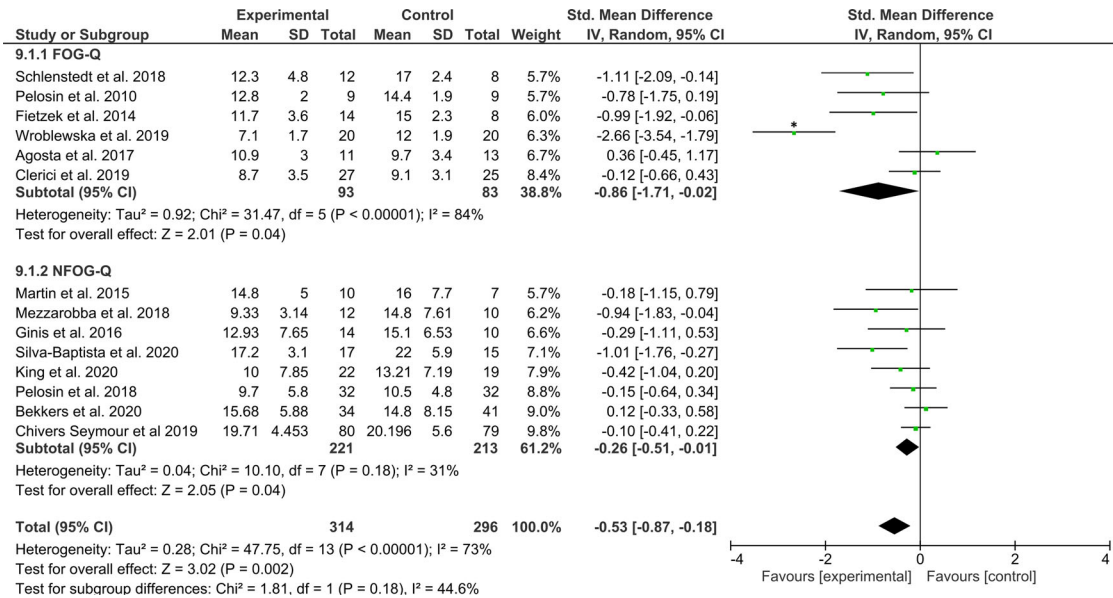

Fig. 5 Secondary analysis comparing the effect of all types of training/exercise against active control groups. The studies are arranged per weight of the study as based on sample size.

A total of 18 studies compared the intervention of interest to an active control of similar exposure. Synthesizing their combined effects across $N_{intervention} = 333$ and $N_{control} = 332$ revealed a favorable moderate effect for the intervention groups ($Z = 3.03$, $p = 0.002$, ES $= -0.30$ [-0.49, -0.11]) with acceptable statistical heterogeneity ($\chi^2(17) = 24.25$, $p = 0.11$, $I^2 = 30\%$) and no significant outliers. The forest plot is presented in Fig. 5.

Not all studies included only patients with FOG. As such, a secondary analysis was performed on 14 studies that only enrolled PD patients who experienced FOG ($N_{intervention} = 314$ and $N_{control} = 296$). The test for overall effect showed a significant moderate to large effect favoring the interventions compared to any type of control group ($Z = 3.02$, $p = 0.002$, ES $= -0.53$ [-0.87, -0.18]), but with high statistical heterogeneity across study effects ($\chi^2(13) = 47.75$, $p < 0.00001$, $I^2 = 73\%$) (Fig. 6). A sensitivity analysis revealed

Fig. 6 Secondary analysis comparing all types of training/exercise interventions against any type of control group in studies enrolling the only PD with FOG (i.e., freezers). The studies are arranged per weight of the study as based on sample size; *significant outliers removed in the sensitivity analysis.

that the heterogeneity in the overall effect was mainly driven by Wroblewska et al., exclusion of which resulted in ($\chi^2(12) = 19.31$, $p = 0.08$, $I^2 = 38\%$) and a significant moderate effect still favoring the intervention ($Z = 2.71$, $p = 0.007$, ES $= -0.32$ [-0.55, -0.09]). Of these studies, ten designed the intervention to specifically target FOG or its underlying correlates (category A and B)[40,41,43–46,48,49,53,55]. Wroblewska et al. also mention that the Nordic walking training was delivered specifically for FOG, though the authors provided no hypothesis on how the intervention would indeed reduce FOG[54]. The home-based smartphone delivered automated feedback training intervention trialed by Ginis et al. was aimed at improving gait in PD, including, but not specific to FOG[70]. Similarly, the virtual-reality treadmill training studied by Bekkers et al. signified a secondary analysis on FOG of a trial that was originally designed to reduce falls[89]. Chivers-Seymour et al. trialed the PD-SAFE program designed for fall prevention in PD and not for FOG in particular[82]. Exploratory removal of these three studies, as well as the significant outlier study of Wroblewska et al.[54], revealed that the overall effects of the remaining studies specific to FOG ($N_{intervention} = 166$, $N_{control} = 146$) remained statistically significant ($Z = 3.08$, $p = 0.002$, ES $= -0.46$ [-0.76, -0.17]), with acceptable statistical heterogeneity across study effects ($\chi^2(9) = 13.8$, $p = 0.13$, $I^2 = 35\%$). Importantly, this indicates that training-based interventions specifically targeting FOG are effective in reducing subjective FOG severity in PD freezers.

For the second research question of this review, we categorized the interventions according to their relevance to FOG (Box 1). The interventions were split into those that specifically aimed to alleviate FOG episodes or circumvent FOG-provoking situations (12 studies in meta-analysis), those that targeted the underlying motor and/or non-motor correlates of FOG (13 studies); and finally, those that offered generic exercise or physical therapy for other health benefits, except FOG (16 studies).

The first category of studies, which were at least partly directed at the alleviation of FOG episodes ($N_{intervention} = 403$ and $N_{control} = 404$ patients), revealed a significant moderate effect, favoring the interventions ($Z = 3.12$, $p = 0.002$, ES $= -0.35$ [-0.56, -0.13]) with some statistical heterogeneity across study effects ($\chi^2(11) = 20.11$, $p < 0.04$, $I^2 = 45\%$) (Fig. 7). A sensitivity analysis without an outlier[87], revealed that the test for overall effects remained significant ($Z = 2.84$, $p = 0.005$, ES $= -0.24$ [-0.40,

-0.07]) with acceptable levels of statistical heterogeneity across study effects ($\chi^2(10) = 11.57$, $p = 0.31$, $I^2 = 14\%$).

The second category, aimed at the underlying correlates of FOG, including 263 ($N_{intervention}$) and 159 ($N_{control}$) patients revealed a significant moderate effect favouring the intervention ($Z = 3.28$, $p = 0.001$, ES $= -0.59$ [-0.95, -0.24]), but again with high statistical heterogeneity across study effects ($\chi^2(12) = 43.24$, $p < 0.0001$, $I^2 = 72\%$), which was mainly driven by the study of Wroblewska et al.[54] (Fig. 8). A sensitivity analysis without this outlier showed that the test for overall effects remained significant in favor of the intervention ($Z = 3.27$, $p = 0.001$, ES $= -0.40$ [-0.64, -0.16]) with acceptable levels of statistical heterogeneity across study effects ($\chi^2(11) = 17.18$, $p = 0.10$, $I^2 = 36\%$).

The third and final subtype analysis on general exercise interventions across a total of $N_{intervention} = 267$ and $N_{control} = 242$ patients revealed a significant, yet small effect of exercise compared to any type of control group ($Z = 2.03$, $p = 0.04$, ES $= -0.20$ [-0.39, -0.01]), with little statistical heterogeneity across study effects ($\chi^2(15) = 16.94$, $p = 0.32$, $I^2 = 11\%$) (Fig. 9). The funnel plot, however, indicated that the effects were likely driven by a significant outlier[72] (Supplementary Data 1). The overall small effect was indeed no longer significant after removal of this one study ($Z = 1.57$, $p = 0.12$, ES $= -0.14$ [-0.32, 0.04], $\chi^2(14) = 9.42$, $p = 0.80$, $I^2 = 0\%$).

To further explore whether training/exercise interventions continue to reduce self-reported FOG beyond the intervention period, a final secondary retention analysis was performed on data from a subgroup of 15 studies that assessed the (N)FOG-Q at follow-up (1–6 months) after a period without training. Figure 10 shows no favorable effect of the interventions for reducing subjective FOG severity at follow-up ($N_{intervention} = 257$, $N_{control} = 263$, $Z = 1.64$, $p = 0.10$, ES $= -0.16$ [-0.36, 0.03]) with little heterogeneity across study effects ($\chi^2(14) = 16.71$, $p = 0.27$, $I^2 = 16\%$). Inspection of the funnel-plot showed that one of these studies[51] was a significant outlier, see Supplementary Data 1. Removal of this study lowered the statistical heterogeneity to ($\chi^2(13) = 10.32$, $p = 0.67$, $I^2 = 0\%$), while the overall small effect remained non-significant ($Z = 0.91$, $p = 0.36$, ES $= -0.08$ [-0.27, 0.10]). Moreover, three of the included studies[61,75,90] were of the generic exercise category and thus not designed to specifically counter FOG. Removal of these three studies along with the one outlier[51] made no
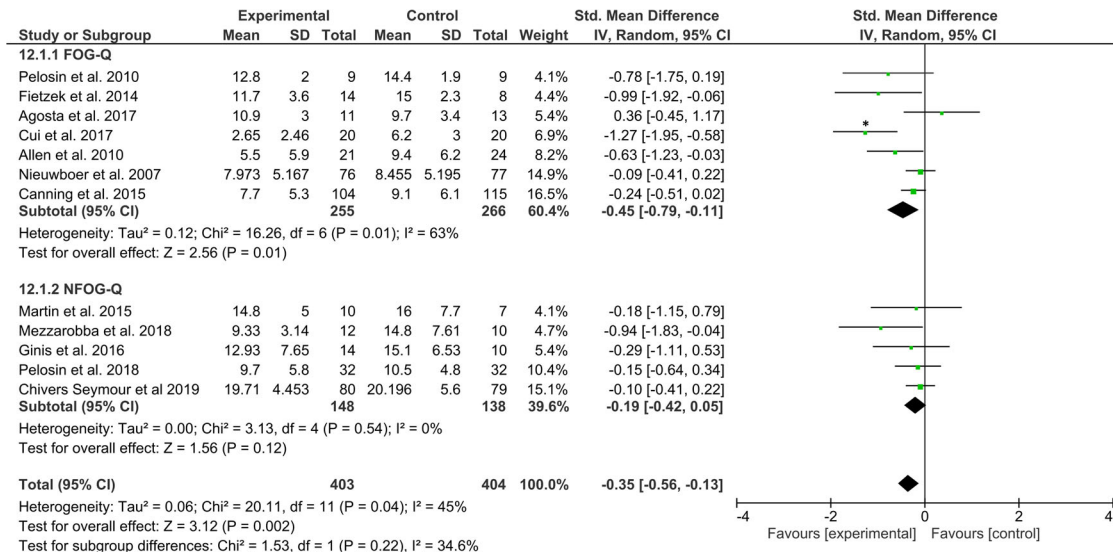
Fig. 7 Secondary analysis comparing effects of category A interventions aimed directly at the alleviation of FOG episodes or FOG-provoking triggers against any type of control group. The studies are arranged per weight of the study as based on sample size; *significant outliers removed in the sensitivity analysis.
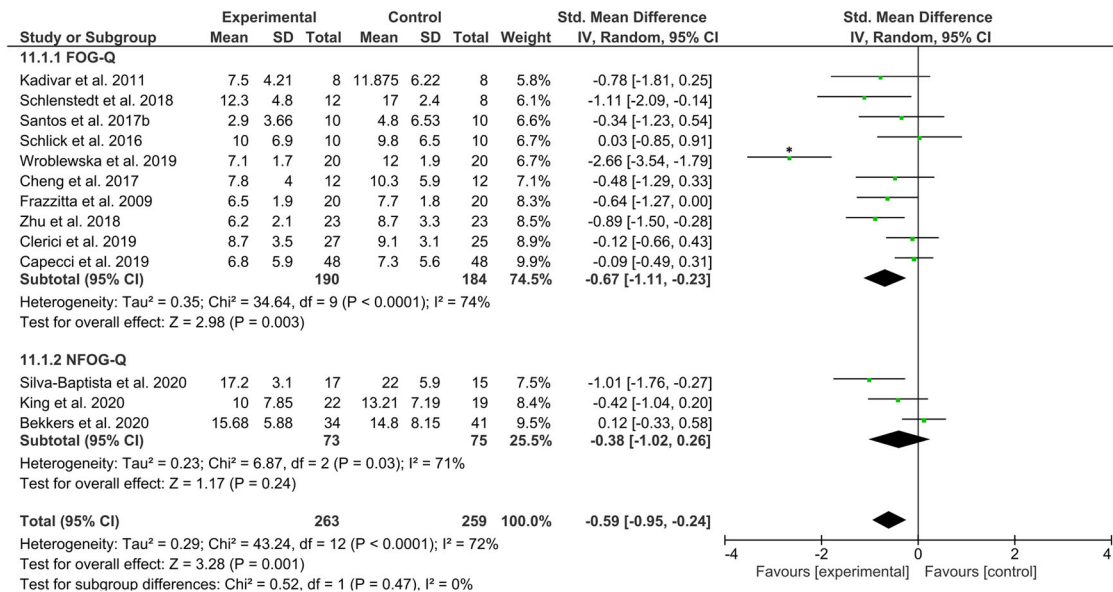

Fig. 8 Secondary analysis comparing effects of category B interventions aimed at the underlying correlates of FOG against any type of control group. The studies are arranged per weight of the study as based on sample size; *significant outliers removed in the sensitivity analysis.

difference to the overall effect size, which was still too small to reach statistical significance ($Z = 0.88$, $p = 0.38$, ES $= -0.09$ [-0.30, 0.11]) with little heterogeneity ($\chi^2(10) = 9.48$, $p = 0.49$, $I^2 = 0\%$).

DISCUSSION

This systematic review with meta-analysis assessed the published effects of exercise- and training-based interventions to reduce FOG-severity in PD. The primary meta-analysis, including 41 studies and 1838 PD patients, revealed a favorable and small-to-moderate effect of a wide variety of training modes for reducing subjective FOG-severity compared to any type of control condition, and this with acceptable levels of statistical heterogeneity across study effects. Moreover, exercise/training of any type proved more beneficial for reducing self-reported FOG severity than no intervention, or an active control condition. Importantly, training/

exercise offered to only freezers, a notoriously more severely affected subgroup of PD on both motor and non-motor outcomes, proved also more beneficial than any type of control intervention, constituting the most salient finding of this study. However, no retention effects were found, indicating that the effects of exercise/training were not sustained beyond the intervention period. Finally, our subgroup analysis revealed that interventions aimed directly at reducing FOG and those aimed at the underlying correlates of FOG were both successful in reducing self-reported severity, whilst generic exercises were not.

The outcomes of the primary analysis assessing the (N)FOG-Q scores after any type of exercise/training compared to any type of control condition revealed a modestly positive and highly significant effect of the various exercise/training interventions for reducing subjective FOG in PD. These effects also remained significant and in favor of the exercise/training interventions when
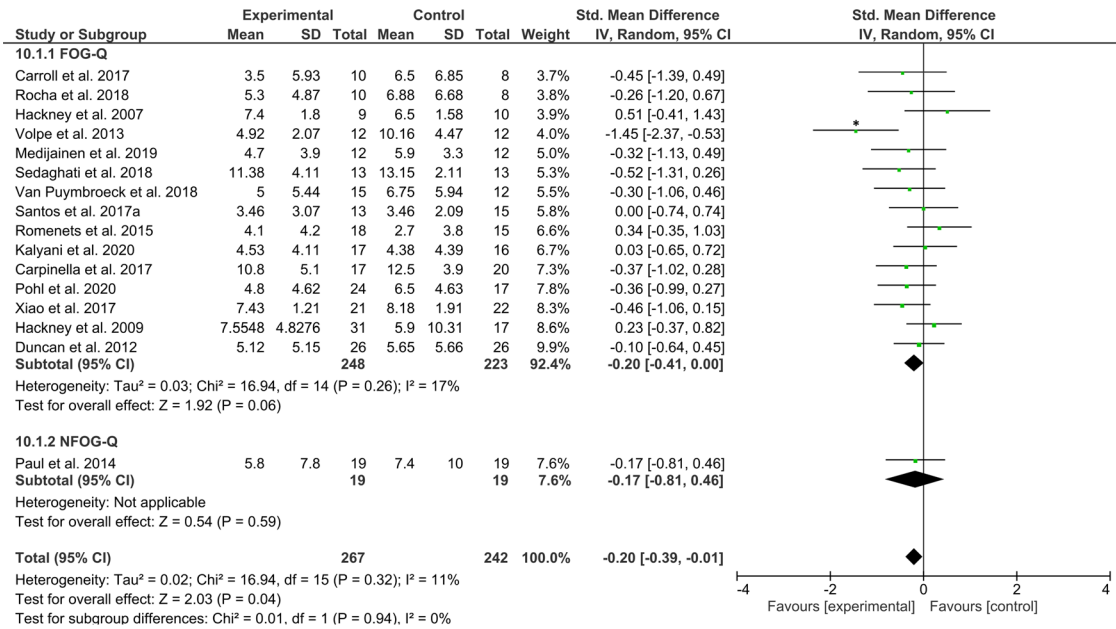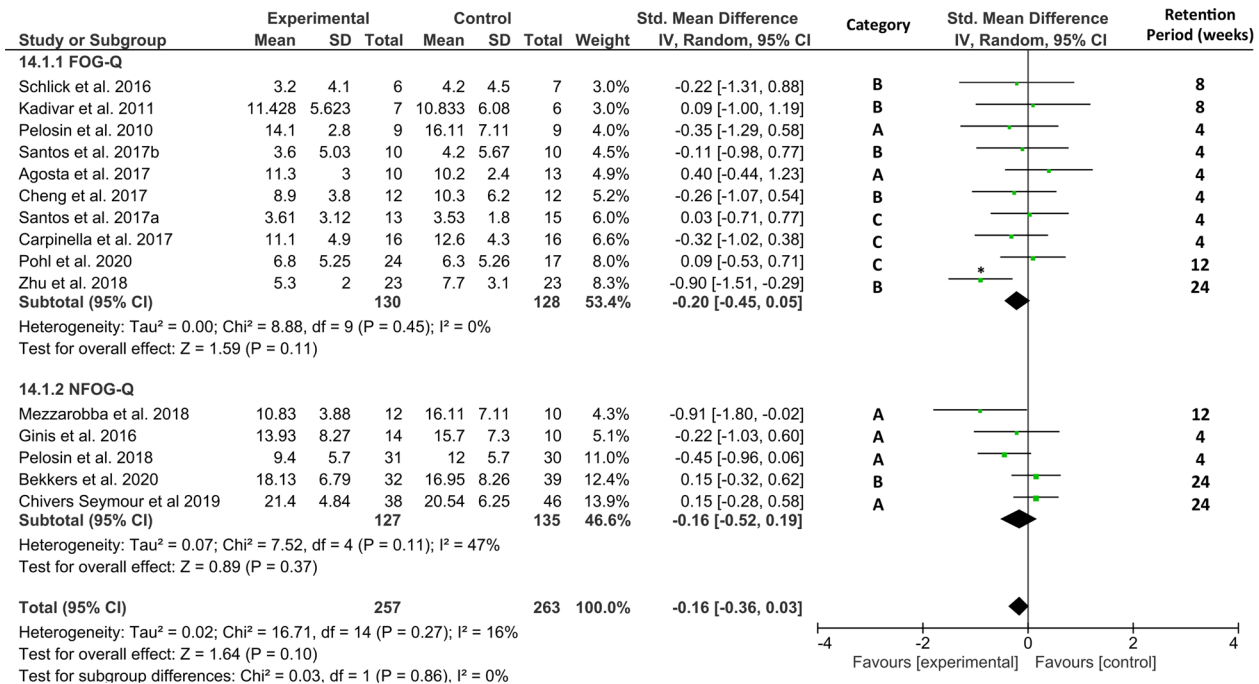
Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
10.1.1 FOG-Q								
Carroll et al. 2017	3.5	5.93	10	6.5	6.85	8	3.7%	-0.45 [-1.39, 0.49]
Rocha et al. 2018	5.3	4.87	10	6.88	6.68	8	3.8%	-0.26 [-1.20, 0.67]
Hackney et al. 2007	7.4	1.8	9	6.5	1.58	10	3.9%	0.51 [-0.41, 1.43]
Volpe et al. 2013	4.92	2.07	12	10.16	4.47	12	4.0%	-1.45 [-2.37, -0.53]
Medijainen et al. 2019	4.7	3.9	12	5.9	3.3	12	5.0%	-0.32 [-1.13, 0.49]
Sedaghati et al. 2018	11.38	4.11	13	13.15	2.11	13	5.3%	-0.52 [-1.31, 0.26]
Van Puymbroeck et al. 2018	5	5.44	15	6.75	5.94	12	5.5%	-0.30 [-1.06, 0.46]
Santos et al. 2017a	3.46	3.07	13	3.46	2.09	15	5.8%	0.00 [-0.74, 0.74]
Romenets et al. 2015	4.1	4.2	18	2.7	3.8	15	6.6%	0.34 [-0.35, 1.03]
Kalyani et al. 2020	4.53	4.11	17	4.38	4.39	16	6.7%	0.03 [-0.65, 0.72]
Carpinella et al. 2017	10.8	5.1	17	12.5	3.9	20	7.3%	-0.37 [-1.02, 0.28]
Pohl et al. 2020	4.8	4.62	24	6.5	4.63	17	7.8%	-0.36 [-0.99, 0.27]
Xiao et al. 2017	7.43	1.21	21	8.18	1.91	22	8.3%	-0.46 [-1.06, 0.15]
Hackney et al. 2009	7.5548	4.8276	31	5.9	10.31	17	8.6%	0.23 [-0.37, 0.82]
Duncan et al. 2012	5.12	5.15	26	5.65	5.66	26	9.9%	-0.10 [-0.64, 0.45]
Subtotal (95% CI)			248			223	92.4%	-0.20 [-0.41, 0.00]

Heterogeneity: Tau² = 0.03; Chi² = 16.94, df = 14 (P = 0.26); I² = 17%
Test for overall effect: Z = 1.92 (P = 0.06)

10.1.2 NFOG-Q								
Paul et al. 2014	5.8	7.8	19	7.4	10	19	7.6%	-0.17 [-0.81, 0.46]
Subtotal (95% CI)			19			19	7.6%	-0.17 [-0.81, 0.46]

Heterogeneity: Not applicable
Test for overall effect: Z = 0.54 (P = 0.59)

| **Total (95% CI)** | | | 267 | | | 242 | 100.0% | -0.20 [-0.39, -0.01] |

Heterogeneity: Tau² = 0.02; Chi² = 16.94, df = 15 (P = 0.32); I² = 11%
Test for overall effect: Z = 2.03 (P = 0.04)
Test for subgroup differences: Chi² = 0.01, df = 1 (P = 0.94), I² = 0%

Fig. 9 Secondary analysis comparing category C "generic exercise" interventions against any type of control group. The studies are arranged per weight of the study as based on sample size; *significant outliers removed in the sensitivity analysis.

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI	Category	Retention Period (weeks)
14.1.1 FOG-Q										
Schlick et al. 2016	3.2	4.1	6	4.2	4.5	7	3.0%	-0.22 [-1.31, 0.88]	B	8
Kadivar et al. 2011	11.428	5.623	7	10.833	6.08	6	3.0%	0.09 [-1.00, 1.19]	B	8
Pelosin et al. 2010	14.1	2.8	9	16.11	7.11	9	4.0%	-0.35 [-1.29, 0.58]	A	4
Santos et al. 2017b	3.6	5.03	10	4.2	5.67	10	4.5%	-0.11 [-0.98, 0.77]	B	4
Agosta et al. 2017	11.3	3	10	10.2	2.4	13	4.9%	0.40 [-0.44, 1.23]	A	4
Cheng et al. 2017	8.9	3.8	12	10.3	6.2	12	5.2%	-0.26 [-1.07, 0.54]	B	4
Santos et al. 2017a	3.61	3.12	13	3.53	1.8	15	6.0%	0.03 [-0.71, 0.77]	C	4
Carpinella et al. 2017	11.1	4.9	16	12.6	4.3	16	6.6%	-0.32 [-1.02, 0.38]	C	4
Pohl et al. 2020	6.8	5.25	24	6.3	5.26	17	8.0%	0.09 [-0.53, 0.71]	C	12
Zhu et al. 2018	5.3	2	23	7.7	3.1	23	8.3%	-0.90 [-1.51, -0.29]	B	24
Subtotal (95% CI)			130			128	53.4%	-0.20 [-0.45, 0.05]		

Heterogeneity: Tau² = 0.00; Chi² = 8.88, df = 9 (P = 0.45); I² = 0%
Test for overall effect: Z = 1.59 (P = 0.11)

14.1.2 NFOG-Q										
Mezzarobba et al. 2018	10.83	3.88	12	16.11	7.11	10	4.3%	-0.91 [-1.80, -0.02]	A	12
Ginis et al. 2016	13.93	8.27	14	15.7	7.3	10	5.1%	-0.22 [-1.03, 0.60]	A	4
Pelosin et al. 2018	9.4	5.7	31	12	5.7	30	11.0%	-0.45 [-0.96, 0.06]	A	4
Bekkers et al. 2020	18.13	6.79	32	16.95	8.26	39	12.4%	0.15 [-0.32, 0.62]	B	24
Chivers Seymour et al 2019	21.4	4.84	38	20.54	6.25	46	13.9%	0.15 [-0.28, 0.58]	A	24
Subtotal (95% CI)			127			135	46.6%	-0.16 [-0.52, 0.19]		

Heterogeneity: Tau² = 0.07; Chi² = 7.52, df = 4 (P = 0.11); I² = 47%
Test for overall effect: Z = 0.89 (P = 0.37)

| **Total (95% CI)** | | | 257 | | | 263 | 100.0% | -0.16 [-0.36, 0.03] | | |

Heterogeneity: Tau² = 0.02; Chi² = 16.71, df = 14 (P = 0.27); I² = 16%
Test for overall effect: Z = 1.64 (P = 0.10)
Test for subgroup differences: Chi² = 0.03, df = 1 (P = 0.86), I² = 0%

Fig. 10 Secondary analysis comparing retention effects of all types of training/exercise interventions against any type of control group. The studies are arranged per weight of the study as based on sample size; *significant outliers removed in the sensitivity analysis.

compared to either passive- or active control conditions. These findings thereby corroborate the outcomes of a recent meta-analysis[37], whilst including data from additional studies increasing the robustness of these prior findings. More importantly, they indicate that diverse modes of exercise and/or training can be offered to PD to reduce the impact of FOG. Careful interpretation is warranted, however, given that many interventions were not directly aimed at reducing FOG, or FOG-related deficits, and included non-freezers. Item 1 of the NFOG-Q enables the exclusion of non-freezers, whom would all have scored zero on this outcome. The "original" FOG-Q, however, does not exclude

non-freezers and contains two non-FOG-specific items. As such, non-freezers can score a maximum of 8 points on this scale[64]. Given that several studies included non-freezers, it cannot be established with certainty that the findings were truly FOG-specific. For this reason, we also conducted a meta-analysis in freezers only, underscoring that training-based interventions were indeed effective in this disease-burdened target population. Although some of these studies utilized the original FOG-Q containing two non-FOG-specific items, the outcome of this sub-analysis provides even stronger evidence of positive training effects on FOG in PD.
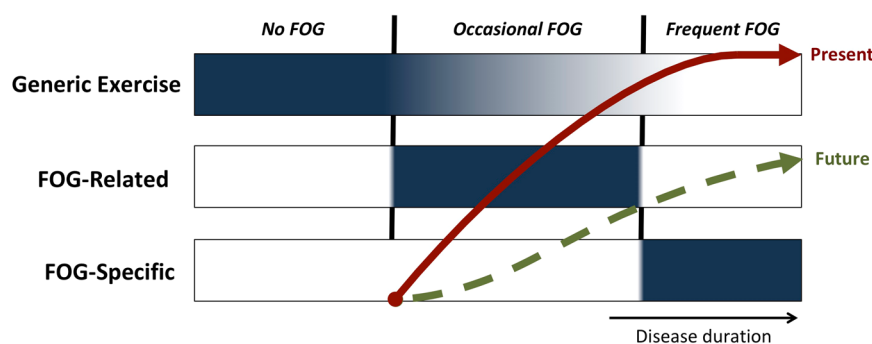
Fig. 11 Theoretical framework for selecting the type of exercise/training in PD using a simple selection criteria, namely the self-reported frequency of FOG. Blue shaded areas indicate when this category of intervention is recommended. The red arrow represents the current progression of worsening FOG severity. The green dotted arrow represents the hypothesized attenuated progression of FOG severity that may be achieved by following this framework. Note that the intensity of generic exercise will likely need to be reduced when FOG-related interventions are offered due to time, energy, and resource constraints. There will also be an increasing need for supervised training by a therapist as the disease progresses and FOG becomes more regular.

Another pertinent question we aimed to address was: which exercise/training intervention is best for FOG? This question remains difficult to answer, given the large variety of interventions trialed to date and because only a few studies directly compared the effects of different modes of exercise/training[55,62,67,68,72,79,80]. Until now, no consensus exists on a gold-standard intervention to which all other interventions can be compared, reflecting the current arbitrary clinical approach. It would therefore have been, and probably still is, more pressing to compare exercise/training against a passive or sham-control condition to test whether the intervention of interest is indeed effective for reducing FOG, before comparing it against another intervention.

Despite these drawbacks, we found that generic exercises do not contribute to the alleviation of FOG, in contrast to FOG-specific and FOG-relevant training. Surprisingly, FOG-relevant (category B) training was even more effective than FOG-specific (category A) training. This is unlikely explained by demographic or clinical differences. The average age (category A = 68.7; B = 69.9), disease duration (A = 9.4, B = 8.5 years), UDPRS-III (A = 28.4; B = 28.8), and (N)FOG-Q scores (A = 12.3; B = 11.4) at baseline did not appear to be different between these groups as based on available data from studies included in the subgroup meta-analysis (Supplementary Table 1). The surplus value of FOG-relevant exercise may be explained by the fact that such training plays into the notion that FOG-episodes are triggered by context-dependent dysfunctional neural information from various (compensatory) neural regions, which ultimately influences the locomotor network and causes transient gait disruptions[13,14,17,19]. As such, FOG-relevant exercise may increase the robustness of this compensatory reserve and prevent or postpone the actual emergence of FOG. In line with other work in neurorehabilitation for PD, training-related alterations are most likely to strengthen neuroplasticity through modulating compensatory circuits rather than changing the affected regions, such as the posterior putamen[91]. However, as the precise etiology of FOG remains elusive, i.e., whether it results from a temporary disruption (or overburdening) in the compensatory networks that allows for FOG to emerge, a dysfunction in the core systems that triggers the episodes, or a combination, it is probably most adequate at present to target both. This is supported by the present finding that both the FOG-specific (targeted triggers mostly) and FOG-relevant (targeted determinants mostly) type interventions seemed efficacious for reducing subjective FOG severity.

What is more, our findings provide some new insights for translation to the clinical field. We propose a new framework for selecting the type of exercise/training for managing FOG using a very simple selection criterion, namely the frequency of FOG experienced (Fig. 11).

We suggest, together with others[92,93], that conventional physical therapy and generic exercises should be made available to all PD patients. Generic exercise is best started early and maintained for as long as possible at a moderate to vigorous rate as this has been shown to enhance physical and mental fitness[92], improve sleep quality[94], and possibly even slow down motor symptom progression[93,95]. There is also indication that exercise increases striatal dopamine release in the anterior striatum[96], and as such, exercise may help facilitate compensatory gait processing and subsequently delay the onset of FOG (Fig. 1), though this remains speculative at present. Early engagement in exercise may also lead to better adherence to long-term physical therapy interventions later in the disease[97]. Our review, as well as those of others[37], have shown the benefit of a large variety of exercises (e.g., swimming, cycling, dancing, tai-chi, yoga, etc.) which can be offered according to the preferences (i.e., for optimal enjoyment) and abilities of each patient. Early in the disease when the risk of falling or other adverse events is relatively low, such generic exercises would not require constant supervision by trained staff, making it more widely accessible. Certain exercise programs, such as Tango dancing and Nordic walking, may also address some of the motor (e.g., gait, balance) and non-motor (e.g., cognitive) correlates of FOG and facilitate rhythmic movements via external cueing that could prove beneficial for pwPD who already experience FOG[54,77,98].

Figure 11 illustrates that we further suggest that PD patients who present with occasional FOG (e.g., <1× week mostly in the dopaminergic OFF state), will likely benefit most from interventions targeting the underlying motor and non-motor correlates of FOG, making compensatory brain circuits more resilient against gait breakdown. Based on converging evidence, therapists may soon be able to predict which non-freezing patients have an increased risk for developing FOG in the future. Studies, so far, show that patients with mainly axial signs, gait difficulties, and repetitive movement breakdown are at risk for conversion to FOG[2,99]. These risk groups are likely to benefit from FOG-relevant interventions to reduce the impact of FOG in the future and perhaps even delay FOG onset. The most likely candidate training modes involve balance training[55,100], turning practice[52], cognitive training[50], and combined motor-cognitive (i.e., dual-task) training[43,47,48]. Group-based classes could be considered whenever feasible for such interventions, as in one study these were shown to be most effective for reducing FOG[79]. The authors propose that the group-based classes may have challenged the patients more in terms of dividing attention and cognitive function than the individual delivery modes, which may thereby have resulted in these extra gains[79]. Indeed, recent studies showed that group

programs combining such exercise also led to modest improvements of FOG[43,48].

In the third column of our model (Fig. 11), we show that PD patients who are troubled by regular FOG are probably best-offered interventions aimed at directly managing FOG episodes. The most well-studied examples are cueing and cognitive rescue strategies to prevent or alleviate FOG[40,41,56,57]. Other training modes involve providing online gait feedback[70], or preparing patients for FOG-provoking circumstances, such as through action-observation training[44–46,53]. Also, fall-prevention strategies in daily complex environments with the ultimate goal of helping patients cope with the FOG-related falls are part of this menu[74,82,86,101]. Given that anxiety can exacerbate FOG[11,102], training behavioral strategies to better deal with the stress in anticipation of upcoming FOG, is also a viable FOG-specific intervention[103], though in this systematic review we could not locate a study specifically trialing such a strategy. Therapists should closely supervise the delivery of FOG-related interventions to ensure correct performance, safety, and adjust difficulty levels over time for each patient[38]. To promote translation to everyday life, supervised home-based delivery may need to be considered[37,42].

We speculate, as shown in Fig. 11, that it may be possible to shift the current projected evolution of FOG, as indicated by the red line, to a slower progression, as shown by the green dotted line. This possibility may be strengthened by capitalizing on the accumulative effect of FOG-relevant and FOG-specific interventions together. Moreover, offering a variety of FOG-related interventions might help keep patients motivated and their frontal attention circuits in an early motor learning state and thereby trained to better control motor operations. In further analogy to motor learning, a combination of aerobic exercise and motor training was found to have a synergistic effect, enhancing neuroplasticity in motor-related compensatory circuits in PD[104]. Careful selection of linking a specific mode of FOG-relevant training with a FOG-specific management approach may be also be favored to address the heterogeneity of FOG and its differing phenotypical manifestations[105]. It thereby remains imperative to keep improving our understanding of the determinants of FOG so that these can be targeted with future behavioral interventions and verify their impact experimentally.

However, the current study also revealed the lack of retention effects. Based on 15 studies, it became apparent that the effects of physical activity waned-off within 1–6 months after ceasing the intervention. This finding contradicts a recent meta-analysis, done on only eight studies, showing sustained retention effects after physical therapy[37]. The same eight studies were included in the present analysis, as were an additional seven. Also, the study of Zhu et al.[51], which in our analysis was considered an outlier, was included in the study by Cosentino et al.[37]. Regardless, the lack of significant retention highlights the importance of keeping freezers engaged in exercise/training in the long-term, which is not self-evident in this population[38,97]. FOG itself, as well as prior falls and fear of falling, are in fact key predictors of poor adherence to exercise in PD[97]. Common non-motor symptoms, including anxiety, depression, cognitive decline, and pain are also predictive of poor adherence in PD[97], highlighting the need for multi-disciplinary care to achieve optimal disease management in freezers[106]. Freezers in particular may thus require extra motivation and follow-up. The outcomes of our review, fortunately, indicate that various training modes may be effective for FOG, offering a wide choice for patients. Adherence in the short-term at least was high (70–100%) for the many interventions reviewed here. This suggests that most people with PD are willing and able to engage in physical therapy interventions in the context of clinical research projects.

With the recent development of technology, another approach is to provide freezers with continuous intelligent cueing and feedback for gait. It is now increasingly possible to personalize cueing and make it fit directly to the motor output of the individual[107,108]. In addition, relatively low cost, lightweight, and unobtrusive Smartphone-based solutions will be leveraged in the near future to deliver cueing essentially at any time and in any setting. However, major impediments to the wide clinical use of wearable systems are that the development of valid algorithm to detect the freezing episodes online and accurately is still challenging.

There are several critical points that should be kept in mind when interpreting the outcomes of the present meta-analyses. First, as also previously noted by others[109], the quality of reporting was fair at best for the majority of studies (Table 2) and did often not comply with the CONsolidated Standards of Reporting Trials (CONSORT) guidelines (consort-statement.org). Secondly, the dosage of the intervention varied a lot between the included studies. This difference was not taken into account in the present meta-analysis. Future studies could provide further insight into the association between treatment dosage and efficacy by means of a meta-regression. Thirdly, a different categorization of studies by intervention types would likely lead to different results in the subgroup analysis. Importantly, however, the categorization of studies in the present review was determined on an a priori basis and conducted by three independent assessors, thereby preventing confirmation bias. Of note is that one study on Nordic walking targeted FOG as the primary outcome[54] and was therefore included in the "FOG-relevant" category, despite offering a conventional exercise type of intervention. Statistically, this study was considered an outlier, and its outcomes were thus not included in the sensitivity analyses. Lastly, only 11 studies defined FOG severity as the primary outcome[41,43–52]. Though, two of these studies investigated an intervention that was not specifically designed to target FOG[47,51]. A reason for not choosing FOG as a primary outcome may have been that FOG severity is notoriously difficult to assess given its unpredictable and episodic nature[1]. Patients with PD are prone to performance bias whereby they tend to switch to goal-directed gait control when being observed. This tendency has particularly negative effects for assessing FOG, because episodes occur less frequently in research or clinical settings[110]. Most studies to date, therefore, included subjective questionnaires to capture FOG severity, which are quick and easy to obtain. However, these scales can be biased by recollection errors, whereby patients may find it difficult to rate the severity of their symptoms in hindsight and scores may be influenced by a single troubling episode. Importantly, given that blinding of participants is difficult to achieve in exercise/training interventions, subjective ratings may be susceptible to placebo effects. Finally, our recent work has shown that the minimal detectable change of the NFOG-Q is 9.95 points, which constitutes about a third of the full range of scores[111]. Taken together, we do not recommend future studies rely solely on the (N)FOG-Q as an outcome of FOG severity.

Efforts are being made to develop responsive measures of FOG severity for use as outcomes in clinical trials, for instance by means of the semi-objective FOG score[69] or by calculating the percentage of time spent frozen as rated from video-recordings of standardized FOG provoking walking tasks[50]. Although it is projected that the field will move towards fully objective at-home ratings of FOG severity as measured with inertial sensors, the limited accuracy of these systems to date for detecting FOG precludes their current use as a primary outcome[112]. We, therefore, recommend future phase-II and phase-III trials to be based on percentage time frozen obtained from video recordings of standardized walking tasks[50] as the primary outcomes of choice for FOG severity[88]. Software to annotate the video recordings for FOG can be downloaded for free from morangi-lat.com[88]. The standardized protocol should consist of a substantial number of FOG-provoking tasks, such as turning on

the spot[69,71,113]. The percentage time frozen can be rated by independent investigators who are kept blinded to group allocation[41,50,55]. Overall, the development of standardization procedures for assessment protocols and sensor-based methods is the precondition for further advancing trials that address novel therapeutic options for FOG.

CONCLUSION

This systematic review with meta-analysis revealed a small-to-moderate effect size favoring various targeted training modes, but not generic exercise, for reducing the subjective impact of FOG in PD. Large-scale RCT's investigating training interventions specifically aimed at reducing FOG using adequate outcomes are still urgently needed to further optimize the multidisciplinary management required for this common and disabling symptom in PD.

DATA AVAILABILITY

The data that support the findings of this study are available within the paper and its Supplementary Information files.

REFERENCES

1. Nutt, J. G. et al. Freezing of gait: moving forward on a mysterious clinical phenomenon. *Lancet Neurol.* **10**, 734–744 (2011).
2. Giladi, N. et al. Freezing of gait in PD: prospective assessment in the DATATOP cohort. *Neurology* **56**, 1712–1721 (2001).
3. Pelicioni, P. H. S., Menant, J. C., Latt, M. D. & Lord, S. R. Falls in Parkinson's disease subtypes: risk factors, locations and circumstances. *Int. J. Environ. Res. Public Health* **16**, 2216 (2019).
4. Schaafsma, J. D. et al. Characterization of freezing of gait subtypes and the response of each to levodopa in Parkinson's disease. *Eur. J. Neurol.* **10**, 391–398 (2003).
5. Spildooren, J. et al. Freezing of gait in Parkinson's disease: the impact of dual-tasking and turning. *Mov. Disord.* **25**, 2563–2570 (2010).
6. Lord, S. R. et al. Freezing of gait in people with Parkinson's Disease: nature, occurrence, and risk factors. *J. Parkinsons Dis.* **10**, 631–640 (2020).
7. Snijders, A. H., Haaxma, C. A., Hagen, Y. J., Munneke, M. & Bloem, B. R. Freezer or non-freezer: clinical assessment of freezing of gait. *Parkinsonism Relat. Disord.* **18**, 149–154 (2012).
8. Cowie, D., Limousin, P., Peters, A. & Day, B. L. Insights into the neural control of locomotion from walking through doorways in Parkinson's disease. *Neuropsychologia* **48**, 2750–2757 (2010).
9. Almeida, Q. J. & Lebold, C. A. Freezing of gait in Parkinson's disease: a perceptual cause for a motor impairment? *J. Neurol. Neurosurg. Psychiatr.* **81**, 513–518 (2010).
10. Ehgoetz Martens, K. A., Pieruccini-Faria, F. & Almeida, Q. J. Could sensory mechanisms be a core factor that underlies freezing of gait in Parkinson's disease? *PLoS ONE* **8**, e62602 (2013).
11. Ehgoetz Martens, K. A., Ellard, C. G. & Almeida, Q. J. Does anxiety cause freezing of gait in Parkinson's disease? *PLoS ONE* **9**, e106561 (2014).
12. Perez-Lloret, S. et al. Prevalence, determinants, and effect on quality of life of freezing of gait in Parkinson disease. *JAMA Neurol.* **71**, 884–890 (2014).
13. Lewis, S. J. G. & Barker, R. A. A pathophysiological model of freezing of gait in Parkinson's disease. *Parkinsonism Relat. Disord.* **15**, 333–338 (2009).
14. Lewis, S. J. G. & Shine, J. M. The next step: a common neural mechanism for freezing of gait. *Neuroscientist* **22**, 72–82 (2014).
15. Snijders, A. H. et al. Physiology of freezing of gait. *Ann. Neurol* **80**, 644–659 (2016).
16. Ehgoetz Martens, K. A. et al. The functional network signature of heterogeneity in freezing of gait. *Brain* **141**, 1145–1160 (2018).
17. Vandenbossche, J. et al. Freezing of gait in Parkinson's disease: disturbances in automaticity and control. *Front. Hum. Neurosci.* **6**, 356 (2012).
18. Bohnen, N. I. et al. Extra-nigral pathological conditions are common in Parkinson's disease with freezing of gait: an in vivo positron emission tomography study. *Mov. Disord.* **29**, 1118–1124 (2014).
19. Gilat, M. et al. Freezing of gait: promising avenues for future treatment. *Parkinsonism Relat. Disord* **52**, 7–16 (2018).

20. Poldrack, R. A. et al. The neural correlates of motor skill automaticity. *J. Neurosci.* **25**, 5356–5364 (2005).
21. Wu, T., Hallett, M. & Chan, P. Motor automaticity in Parkinson's disease. *Neurobiol. Dis.* **82**, 226–234 (2015).
22. Gilat, M. et al. Dopamine depletion impairs gait automaticity by altering cortico-striatal and cerebellar processing in Parkinson's disease. *Neuroimage* **152**, 207–220 (2017).
23. Redgrave, P. et al. Goal-directed and habitual control in the basal ganglia: implications for Parkinson's disease. *Nat. Rev. Neurosci.* **11**, 760–772 (2010).
24. Hausdorff, J. M., Schweiger, A., Herman, T., Yogev-Seligmann, G. & Giladi, N. Dual-task decrements in gait: contributing factors among healthy older adults. *J. Gerontol. A Biol. Sci. Med. Sci.* **63**, 1335–1343 (2008).
25. Hausdorff, J. M. et al. Impaired regulation of stride variability in Parkinson's disease subjects with freezing of gait. *Exp. Brain Res.* **149**, 187–194 (2003).
26. Gilat, M. et al. Dysfunctional limbic circuitry underlying freezing of gait in Parkinsonas disease. *Neuroscience* **374**, 119–132 (2018).
27. Vercruysse, S. et al. Effects of deep brain stimulation of the subthalamic nucleus on freezing of gait in Parkinson's disease: a prospective controlled study. *J. Neurol. Neurosurg. Psychiatr.* **85**, 871–877 (2014).
28. Barbe, M. T. et al. Deep brain stimulation for freezing of gait in Parkinson's disease with early motor complications. *Mov. Disord.* **35**, 82–90 (2020).
29. Nonnekes, J. et al. Freezing of gait: a practical approach to management. *Lancet Neurol.* **14**, 768–778 (2015).
30. Schlenstedt, C. et al. Effect of high-frequency subthalamic neurostimulation on gait and freezing of gait in Parkinson's disease: a systematic review and meta-analysis. *Eur. J. Neurol.* **24**, 18–26 (2017).
31. Kalyani, H. H. N. et al. Effects of dance on gait, cognition, and dual-tasking in Parkinson's disease: a systematic review and meta-analysis. *J. Parkinsons Dis.* **9**, 335–349 (2019).
32. Delabary, Santos et al. Effects of dance practice on functional mobility, motor symptoms and quality of life in people with Parkinson's disease: a systematic review with meta-analysis. *Aging Clin. Exp. Res.* **30**, 727–735 (2018).
33. Ni, M., Hazzard, J. B., Signorile, J. F. & Luca, C. Exercise guidelines for gait function in Parkinson's disease: a systematic review and meta-analysis. *Neurorehabil. Neural Repair* **32**, 872–886 (2018).
34. Lötzke, D., Ostermann, T. & Büssing, A. Argentine tango in Parkinson disease-a systematic review and meta-analysis. *BMC Neurol.* **15**, 226–18 (2015).
35. Rutz, D. G. & Benninger, D. H. Physical therapy for freezing of gait and gait impairments in Parkinson disease: a systematic review. *PM R* **12**, 1140–1156 (2020).
36. Delgado-Alvarado, M. et al. Nonpharmacological, nonsurgical treatments for freezing of gait in Parkinsonas disease: a systematic review. *Mov. Disord* **35**, 204–214 (2019).
37. Cosentino, C. et al. Effectiveness of physiotherapy on freezing of gait in Parkinson's disease: a systematic review and meta-analyses. *Mov. Disord.* **35**, 523–536 (2019).
38. Schootemeijer, S. et al. Barriers and motivators to engage in exercise for persons with Parkinson's disease. *J. Parkinsons Dis.* **10**, 1293–1299 (2020).
39. Bramer, W. M., Rethlefsen, M. L., Kleijnen, J. & Franco, O. H. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Syst. Rev.* **6**, 245 (2017).
40. Martin, T., Weatherall, M., Anderson, T. J. & MacAskill, M. R. A randomized controlled feasibility trial of a specific cueing program for falls management in persons with Parkinson disease and freezing of gait. *J. Neurol. Phys. Ther.* **39**, 179–184 (2015).
41. Fietzek, U. M., Schroeteler, F. E., Ziegler, K., Zwosta, J. & Ceballos-Baumann, A. O. Randomized cross-over trial to investigate the efficacy of a two-week physiotherapy programme with repetitive exercises of cueing to reduce the severity of freezing of gait in patients with Parkinson's disease. *Clin. Rehabil.* **28**, 902–911 (2014).
42. Nieuwboer, A. et al. Cueing training in the home improves gait-related mobility in Parkinson's disease: the RESCUE trial. *J. Neurol. Neurosurg. Psychiatr.* **78**, 134–140 (2007).
43. King, L. A. et al. Cognitively challenging agility boot camp program for freezing of gait in Parkinson disease. *Neurorehabil. Neural Repair* **34**, 417–427 (2020).
44. Mezzarobba, S. et al. Action observation plus sonification. a novel therapeutic protocol for Parkinson's patient with freezing of gait. *Front. Neurol.* **8**, 723 (2017).
45. Pelosin, E. et al. Effect of group-based rehabilitation combining action observation with physiotherapy on freezing of gait in Parkinson's disease. *Neural Plast.* **2018**, 4897276–7 (2018).
46. Pelosin, E. et al. Action observation improves freezing of gait in patients with Parkinson's disease. *Neurorehabil. Neural Repair* **24**, 746–752 (2010).
47. Bekkers, E. M. J. et al. Do patients with Parkinson's disease with freezing of gait respond differently than those without to treadmill training augmented by virtual reality? *Neurorehabil. Neural Repair* **34**, 440–449 (2020).

48. Silva-Batista, C. et al. A randomized, controlled trial of exercise for Parkinsonian individuals with freezing of gait. *Mov. Disord* **35**, 1607–1617 (2020).

49. Clerici, I. et al. Land plus aquatic therapy versus land-based rehabilitation alone for the treatment of freezing of gait in Parkinson disease: a randomized controlled trial. *Phys. Ther.* **99**, 591–600 (2019).

50. Walton, C. C. et al. Cognitive training for freezing of gait in Parkinson's disease: a randomized controlled trial. *NPJ Parkinsons Dis.* **4**, 15 (2018).

51. Zhu, Z.-Z. et al. Aquatic obstacle training improves freezing of gait in Parkinson's disease patients: a randomized controlled trial. *Clin. Rehabil.* **32**, 29–36 (2018).

52. Cheng, F.-Y., Yang, Y.-R., Wu, Y.-R., Cheng, S.-J. & Wang, R.-Y. Effects of curved-walking training on curved-walking performance and freezing of gait in individuals with Parkinson's disease: a randomized controlled trial. *Parkinsonism Relat. Disord.* **43**, 20–26 (2017).

53. Agosta, F. et al. Brain plasticity in Parkinson's disease with freezing of gait induced by action observation training. *J. Neurol.* **264**, 88–101 (2017).

54. Wróblewska, A., Gajos, A., Smyczyńska, U. & Bogucki, A. The therapeutic effect of nordic walking on freezing of gait in Parkinson's disease: a pilot study. *Parkinsons Dis.* **2019**, 3846279 (2019).

55. Schlenstedt, C. et al. Moderate frequency resistance and balance training do not improve freezing of gait in Parkinson's disease: a pilot study. *Front. Neurol.* **9**, 1084 (2018).

56. Schlick, C. et al. Visual cues combined with treadmill training to improve gait performance in Parkinson's disease: a pilot randomized controlled trial. *Clin. Rehabil.* **30**, 463–471 (2016).

57. Frazzitta, G., Maestri, R., Uccellini, D., Bertotti, G. & Abelli, P. Rehabilitation treatment of gait in patients with Parkinson's disease with freezing: a comparison between two physical therapy protocols using visual and auditory cues with or without treadmill training. *Mov. Disord.* **24**, 1139–1143 (2009).

58. Kalyani, H. H. et al. Dance improves symptoms, functional mobility and fine manual dexterity in people with Parkinson disease: a quasi-experimental controlled efficacy study. *Eur. J. Phys. Rehabil. Med.* https://doi.org/10.23736/S1973-9087.20.06069-4 (2020).

59. Sedaghati, P., Goudarzian, M., Daneshmandi, H. & Ardjmand, A. Effects of Alexander-based corrective techniques on forward flexed posture, risk of fall, and fear of falling in idiopathic Parkinsonas disease. *Arch. Neurosci.* **5**, e61274 (2018).

60. Van Puymbroeck, M. et al. Functional improvements in Parkinson's disease following a randomized trial of yoga. *Evid. Based Complement Altern. Med.* **2018**, 8516351 (2018).

61. Santos, L. et al. Effects of progressive resistance exercise in akinetic-rigid Parkinson's disease patients: a randomized controlled trial. *Eur. J. Phys. Rehabil. Med.* **53**, 651–663 (2017).

62. Hackney, M. E. & Earhart, G. M. Effects of dance on movement control in Parkinson's disease: a comparison of Argentine tango and American ballroom. *J. Rehabil. Med* **41**, 475–481 (2009).

63. Hackney, M. E., Kantorovich, S., Levin, R. & Earhart, G. M. Effects of tango on functional mobility in Parkinson's disease: a preliminary study. *J. Neurol. Phys. Ther.* **31**, 173–179 (2007).

64. Giladi, N. et al. Construction of freezing of gait questionnaire for patients with Parkinsonism. *Parkinsonism Relat. Disord.* **6**, 165–170 (2000).

65. Nieuwboer, A. et al. Reliability of the new freezing of gait questionnaire: agreement between patients with Parkinson's disease and their carers. *Gait Posture* **30**, 459–463 (2009).

66. Bloem, B. R. et al. Measurement instruments to assess posture, gait, and balance in Parkinsonas disease: critique and recommendations. *Mov Disord* **31**, 1342–1355 (2016).

67. Pacchetti, C. et al. Active music therapy in Parkinson's disease: an integrative method for motor and emotional rehabilitation. *Psychosom. Med.* **62**, 386–393 (2000).

68. Reuter, I. et al. Effects of a flexibility and relaxation programme, walking, and nordic walking on Parkinson's disease. *J. Aging Res.* **2011**, 232473 (2011).

69. Ziegler, K., Schroeteler, F., Ceballos-Baumann, A. O. & Fietzek, U. M. A new rating instrument to assess festination and freezing gait in Parkinsonian patients. *Mov. Disord.* **25**, 1012–1018 (2010).

70. Ginis, P. et al. Feasibility and effects of home-based smartphone-delivered automated feedback training for gait in people with Parkinson's disease: a pilot randomized controlled trial. *Parkinsonism Relat. Disord.* **22**, 28–34 (2016).

71. Mancini, M. et al. The clinical significance of freezing while turning in Parkinson's disease. *Neuroscience* **343**, 222–228 (2017).

72. Volpe, D., Signorini, M., Marchetto, A., Lynch, T. & Morris, M. E. A comparison of Irish set dancing and exercises for people with Parkinson's disease: a phase II feasibility study. *BMC Geriatr.* **13**, 54 (2013).

73. Rocha, N. P., Lorenna, A., McClelland, J. A. & Morris, M. E. Dance therapy for Parkinson's disease: A randomised feasibility trial. *Int. J. Ther. Rehabilitation* **25**, 64–72 (2018).

74. Canning, C. G. et al. Exercise for falls prevention in Parkinson disease: a randomized controlled trial. *Neurology* **84**, 304–312 (2015).

75. Carpinella, I. et al. Wearable sensor-based biofeedback training for balance and gait in Parkinson disease: a pilot randomized controlled trial. *Arch. Phys. Med. Rehabil.* **98**, 622–630.e3 (2017).

76. Duncan, R. P. & Earhart, G. M. Are the effects of community-based dance on Parkinson disease severity, balance, and functional mobility reduced with time? A 2-year prospective pilot study. *J. Alter. Complement Med.* **20**, 757–763 (2014).

77. Duncan, R. P. & Earhart, G. M. Randomized controlled trial of community-based dancing to modify disease progression in Parkinson disease. *Neurorehabil. Neural Repair* **26**, 132–143 (2012).

78. Xiao, C., Kang, Y. & Zhuang, Y. C. Effect of Tai Chi Ball exercise on functional mobility in mild-moderate Parkinson disease. *J. Gerontol. Geriatr. Res.* **6**, 1–3 (2017).

79. King, L. A. et al. Effects of group, individual, and home exercise in persons with Parkinson disease: a randomized clinical trial. *J. Neurol. Phys. Ther.* **39**, 204–212 (2015).

80. Ricciardi, L. et al. Working on asymmetry in Parkinson's disease: randomized, controlled pilot study. *Neurol. Sci* **36**, 1337–1343 (2015).

81. Carroll, L. M., Volpe, D., Morris, M. E., Saunders, J., & Clifford, A. M. Aquatic exercise therapy for people with Parkinson disease: a randomized controlled trial. *Arch. Phys. Med. Rehabil.* **98**, 631–638 (2017).

82. Chivers-Seymour, K. et al. Multicentre, randomised controlled trial of PDSAFE, a physiotherapist-delivered fall prevention programme for people with Parkinson's. *J. Neurol. Neurosurg. Psychiatr* **90**, 774–782 (2019).

83. Kadivar, Z., Corcos, D. M., Foto, J. & Hondzinski, J. M. Effect of step training and rhythmic auditory stimulation on functional performance in Parkinson patients. *Neurorehabil. Neural Repair* **25**, 626–635 (2011).

84. Hubble, R. P., Silburn, P. A., Naughton, G. A. & Cole, M. H. Trunk exercises improve balance in Parkinson disease: a phase II randomized controlled trial. *J. Neurol. Phys. Ther.* **43**, 96–105 (2019).

85. Byl, N., Zhang, W., Coo, S. & Tomizuka, M. Clinical impact of gait training enhanced with visual kinematic biofeedback: patients with Parkinson's disease and patients stable post stroke. *Neuropsychologia* **79**, 332–343 (2015).

86. Ashburn, A. et al. Exercise- and strategy-based physiotherapy-delivered intervention for preventing repeat falls in people with Parkinson's: the PDSAFE RCT. *Health Technol. Assess.* **23**, 1–150 (2019).

87. Cui, L., Yu, Y., Zhi-zhong, Z., Ya-qing, L., Hou, W., Wang, L. & Cheng, R. Rehabilitation effect of cueing strategies on balance function and gait in patients with Parkinson's disease [article in Chinese]. *Chin. J. Contemp. Neurol. Neurosurg.* **17**, 428–433 (2017).

88. Gilat, M. How to annotate freezing of gait from video: a standardized method using open-source software. *J. Parkinsons Dis.* **9**, 821–824 (2019).

89. Mirelman, A. et al. Addition of a non-immersive virtual reality component to treadmill training to reduce fall risk in older adults (V-TIME): a randomised controlled trial. *Lancet* **388**, 1170–1182 (2016).

90. Pohl, P., Wressle, E., Lundin, F., Enthoven, P. & Dizdar, N. Group-based music intervention in Parkinson's disease—findings from a mixed-methods study. *Clin. Rehabil.* **34**, 533–544 (2020).

91. Nackaerts, E. et al. Towards understanding neural network signatures of motor skill learning in Parkinson's disease and healthy aging. *Br. J. Radio.* **27**, 20190071 (2019).

92. Schootemeijer, S., van der Kolk, N. M., Bloem, B. R. & de Vries, N. M. Current perspectives on aerobic exercise in people with Parkinson's disease. *Neurotherapeutics* **75**, 9–16 (2020).

93. Ellis, T. Exercise in Parkinson's disease: are we narrowing in on the essential elements? *Lancet Neurol.* **18**, 982–983 (2019).

94. Amara, A. W. et al. Randomized, controlled trial of exercise on objective and subjective sleep in Parkinson's disease. *Mov. Disord.* **35**, 947–958 (2020).

95. van der Kolk, N. M. et al. Effectiveness of home-based and remotely supervised aerobic exercise in Parkinson's disease: a double-blind, randomised controlled trial. *Lancet Neurol.* **18**, 998–1008 (2019).

96. Sacheli, M. A. et al. Exercise increases caudate dopamine release and ventral striatal activation in Parkinsonas disease. *Mov. Disord* **34**, 1891–1900 (2019).

97. Allen, N. E. et al. Predictors of adherence to a falls prevention exercise program for people with Parkinson's disease. *Mov. Disord. Clin. Pract.* **2**, 395–401 (2015).

98. Rios Romenets, S., Anang, J., Fereshtehnejad, S.-M., Pelletier, A. & Postuma, R. Tango for treatment of motor and non-motor manifestations in Parkinson's disease: a randomized control study. *Complement Ther. Med.* **23**, 175–184 (2015).

99. D'Cruz, N. et al. Repetitive motor control deficits most consistent predictors of conversion to freezing of gait in Parkinson's disease: a prospective cohort study. *J. Parkinsons Dis.* **10**, 559–571 (2020).

100. Santos, L. et al. Effects of supervised slackline training on postural instability, freezing of gait, and falls efficacy in people with Parkinson's disease. *Disabil. Rehabil.* **39**, 1573–1580 (2017).

101. Allen, N. E. et al. The effects of an exercise program on fall risk factors in people with Parkinson's disease: a randomized controlled trial. *Mov. Disord.* **25**, 1217–1225 (2010).

102. Lieberman, A. Are freezing of gait (FOG) and panic related? *J. Neurol. Sci.* **248**, 219–222 (2006).

103. Maslivec, A., Fielding, A., Wilson, M., Norris, M. & Young, W. 'Recoupling" the attentional and motor control of preparatory postural adjustments to overcome freezing of gait in Parkinson"s. *J. Neuroeng. Rehabil.* **17**, 146–13 (2020).

104. Duchesne, C. et al. Influence of aerobic exercise training on the neural correlates of motor learning in Parkinson's disease individuals. *Neuroimage Clin.* **12**, 559–569 (2016).

105. Ehgoetz Martens, K. A. et al. Evidence for subtypes of freezing of gait in Parkinson's disease. *Mov. Disord.* **33**, 1174–1178 (2018).

106. Radder, D. L. M. et al. Multidisciplinary care for people with Parkinson's disease: the new kids on the block! *Expert Rev. Neurother.* **19**, 145–157 (2019).

107. Barthel, C. et al. The laser shoes: a new ambulatory device to alleviate freezing of gait in Parkinson disease. *Neurology* https://doi.org/10.1212/WNL.0000000000004795 (2017).

108. Ginis, P., Nackaerts, E., Nieuwboer, A. & Heremans, E. Cueing for people with Parkinsonas disease with freezing of gait: a narrative review of the state-of-the-art and novel perspectives. *Ann. Phys. Rehabil. Med* **61**, 407–413 (2017).

109. Silva, C. M., Travessa, A. M., Bouça-Machado, R., Caldeira, D. & Ferreira, J. J. Reporting and methodological quality of clinical trials on exercise therapy for Parkinson's disease. *Parkinsonism Relat. Disord.* **69**, 150–156 (2019).

110. Mancini, M. Clinical and methodological challenges for assessing freezing of gait: Future perspectives. *Mov Disord* **34**, 783–790 (2019).

111. Hulzinga, F. et al. The new freezing of gait questionnaire: unsuitable as an outcome in clinical trials? *Mov. Disord. Clin. Pract.* **7**, 199–205 (2020).

112. Pardoel, S., Kofman, J., Nantel, J. & Lemaire, E. D. Wearable-sensor-based detection and prediction of freezing of gait in Parkinsonas disease: a review. *Sensors* **19**, 5141 (2019).

113. van Dijsseldonk, K., Wang, Y., van Wezel, R., Bloem, B. R. & Nonnekes, J. Provoking freezing of gait in clinical practice: turning in place is more effective than stepping in place. *J. Parkinsons Dis.* **8**, 363–365 (2018).

114. Capecci, M. et al. Clinical effects of robot-assisted gait training and treadmill training for Parkinson's disease. A randomized controlled trial. *Ann. Phys. Rehabil. Med.* **62**, 303–312 (2019).

115. Medijainen, K., Pääsuke, M., Lukmann, A. & Taba, P. Versatile guideline-based physiotherapy intervention in groups to improve gait speed in Parkinson's disease patients. *NeuroRehabilitation* **44**, 579–586 (2019).

116. Paul, S. S., Canning, C. G., Song, J., Fung, V. S. C. & Sherrington, C. Leg muscle power is enhanced by training in people with Parkinson's disease: a randomized controlled trial. *Clin. Rehabil.* **28**, 275–288 (2014).

AUTHOR CONTRIBUTIONS

Conceived and designed the study: M.G., P.G., and A.N. Performed systematic search, screening and quality assessment: M.G. and P.G. Third moderator for screening and quality assessment: N.D. and A.N. Extracted meta-data: M.G. and P.G. Conducted the systematic review and meta-analyses: M.G. Wrote first manuscript: M.G. Construction of figures: M.G. Construction of tables: M.G., P.G., D.Z., J.D.V., and F.H. Review and critique of manuscript: P.G., N.D., D.Z., J.D.V., F.H., and A.N. All authors hold accountability for all aspects of the work.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41531-021-00224-4.

Correspondence and requests for materials should be addressed to Moran Gilat.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.