

Supplementary Table 1 Database-specific literature search strategies

Database	Coverage	Search strategy	Restrictions	Records identified
PubMed	From database inception to May 31, 2026	("Interoception"[MeSH Terms] OR interoception[Title/ Abstract] OR interoceptive[Title/ Abstract] OR "interoceptive accuracy"[Title/ Abstract] OR "interoceptive sensibility"[Title/ Abstract] OR "interoceptive awareness"[Title/ Abstract]) AND ("Depressive Disorder, Major"[MeSH Terms] OR "major depressive disorder"[Title/ Abstract] OR MDD[Title/ Abstract] OR "major depression"[Title/ Abstract])	Publication date up to May 31, 2026; no article-type or full-text filters applied at the database-search stage	98

		TS = (interoception OR interoceptive OR	
	From	"interoceptive accuracy" OR "interoceptive	Timespan: all years to May 31,
Web of	database	sensibility" OR "interoceptive awareness" OR	2026; document type and
Science	inception to	"body awareness" OR "bodily awareness")	full-text availability were not
	May 31, 2026	AND TS = ("major depressive disorder" OR	restricted at the database-search
		MDD OR "major depression" OR "depressive	stage
		disorder")	

199

Abbreviations: MDD, major depressive disorder; MeSH, Medical Subject Headings; TS, Topic Search. Synonyms within each concept were combined using OR, whereas the interoception and MDD concept groups were combined using AND.

Supplementary Table 2 Characteristics, key findings or contributions, and principal limitations of the 69 publications included in the final narrative synthesis

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
1	Verdonk et al., 2025	Narrative/current-opinion review	N/A	Frames interoception as a multilevel brain-body process spanning central, peripheral, computational, and clinical domains.	Selective narrative synthesis; transdiagnostic rather than MDD-specific; no formal evidence grading.
2	Mavar et al., 2025	Umbrella review and conceptual model	Published neuroimaging meta-analyses; no pooled participant N	Synthesizes MDD neuroimaging meta-analyses and proposes large-scale network patterns, including SMN-VN reduction, VAN increase, and stage-related DMN differences.	Built on heterogeneous, largely cross-sectional meta-analytic evidence; the stage-related model is not longitudinally established.
3	Davey, 2025	Conceptual review	N/A	Uses active inference to link altered interoceptive predictions and precision weighting with affective, cognitive, and somatic features of depression.	Theory-led account; key computational mechanisms have limited direct experimental testing in MDD.
4	Khalsa et al., 2018	Consensus roadmap	N/A; expert consensus	Defines major interoceptive constructs and measurement challenges and sets a translational research agenda for mental health.	Expert consensus with broad transdiagnostic scope; not a systematic evidence-grading exercise.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
5	Garfinkel et al., 2015	Cross-sectional experimental study	80 healthy adults	Empirically separated interoceptive accuracy, sensibility, and metacognitive awareness, showing that they are related but non-equivalent.	Healthy sample; cardiac tasks and confidence ratings only; task-specific, cross-sectional operationalization.
6	Eggart et al., 2019	Systematic review	6 eligible cross-sectional MDD studies; participant Ns varied	Reduced heartbeat-perception performance was most evident in moderately depressed samples; findings varied with clinical characteristics and task implementation.	Small, heterogeneous evidence base dominated by heartbeat counting; no longitudinal studies and substantial confounding.
7	Eggart et al., 2025	Prospective observational study	87 hospitalized patients with major depression	Lower MAIA-2 Trusting at admission independently predicted suicidal ideation at discharge, with moderate discrimination.	Single-center inpatient cohort; self-report predictor and outcome; observational design without external validation.
8	Shen et al., 2025	Cross-sectional case-control and mediation study	88 patients with depression and 50 healthy controls (N=138)	Lower self-reported interoceptive sensibility was concentrated in depressed patients with alexithymia; alexithymia mediated its association with depressive symptoms.	Cross-sectional mediation; culturally restricted sample; HCT validity and medication/clinical heterogeneity remain concerns.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
9	Paulus et al., 2019	Conceptual review	N/A	Develops an active-inference account of interoceptive psychopathology based on priors, prediction errors, and precision weighting.	Transdiagnostic and theory-led; limited disorder-specific prospective or experimental validation.
10	Barrett et al., 2016	Theoretical article	N/A	Proposes that disrupted allostasis and interoceptive inference can sustain maladaptive bodily predictions in depression.	Theoretical article without direct empirical testing; proposed computational mechanisms are difficult to operationalize.
11	Seth and Friston, 2016	Theoretical review	N/A	Formulates active interoceptive inference as a mechanism linking predictive bodily regulation, emotion, and embodied selfhood.	General theoretical framework rather than an MDD study; limited direct falsification and clinical validation.
12	Zhou et al., 2026	Two-sample Mendelian-randomiz ation and mediation study	Large GWAS summary datasets across multiple somatic diseases and depression; N varied by trait	Across several somatic diseases, reduced left ventral diencephalon volume and related interoceptive-circuit morphology were implicated in pathways to depression.	Depends on genetic-instrument assumptions and summary-level data; morphology is an indirect proxy for interoceptive function.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
13	Baethge et al., 2019	Reporting-quality scale development	N/A; expert-informed scale development	Introduced SANRA, a six-item scale for assessing the reporting quality of narrative reviews.	Not a risk-of-bias or certainty-of-evidence tool; scoring retains reviewer judgment.
14	Quadt et al., 2018	Narrative review	N/A	Reviews afferent pathways, cortical representations, and broader physiological channels supporting interoception across health and disease.	Broad narrative review; not MDD-specific and integrates heterogeneous mechanistic and clinical evidence.
15	Kaiser et al., 2015	Meta-analysis of resting-state connectivity	27 datasets from 25 publications; 556 MDD and 518 healthy controls	Identified reproducible MDD-related dysconnectivity across default-mode, salience/affective, and cognitive-control systems.	Coordinate/seed-based literature with marked analytical and clinical heterogeneity; predominantly cross-sectional.
16	Mulders et al., 2015	Review of resting-state connectivity	36 studies (8 ICA and 28 seed-based studies)	Reported recurrent anterior DMN hyperconnectivity, altered anterior-posterior DMN coupling, and abnormal interactions with salience/executive networks.	Large methodological heterogeneity in seeds, preprocessing, samples, and medication status; qualitative synthesis.
17	Barrett and Simmons, 2015	Conceptual review	N/A	Proposes the Embodied Predictive Interoception Coding (EPIC) model, in which cortical interoceptive predictions guide allostasis and are constrained by ascending bodily signals.	Not depression-specific; model is primarily theoretical and neural mappings remain partly inferential.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
18	Zhu et al., 2012	Cross-sectional resting-state fMRI case-control study	35 first-episode, treatment-naive MDD patients and 35 matched controls	First-episode, treatment-naive MDD showed increased anterior and reduced posterior DMN connectivity, indicating subsystem dissociation.	Small single-center sample; cross-sectional design; results depend on ICA/connectivity choices.
19	Hamilton et al., 2015	Narrative review	N/A	Links depressive rumination to DMN function and its interaction with cognitive-control systems.	Narrative synthesis not focused on interoception; causal direction between network activity and rumination remains unresolved.
20	Wu et al., 2025	Secondary longitudinal structural-MRI analysis	85 remitted depressed outpatients from a psychotherapy trial	Cortical thickness in insular and related regions was associated with depression vulnerability, residual symptoms, and regulatory processes in remitted patients.	Secondary analysis with modest sample; remitted outpatient population; regional associations do not establish mechanism.
21	Horn et al., 2010	Cross-sectional MRS and resting-state fMRI study	18 MDD patients and 22 healthy controls	Lower pgACC glutamate-glutamine and altered pgACC-anterior-insula connectivity were most evident in patients with greater depression severity.	Small sample and severity subgroups; cross-sectional MRS/fMRI; medication and voxel/connectivity constraints.
22	Miller and Raison, 2016	Narrative mechanistic review	N/A	Synthesizes evidence linking immune activation and inflammation with depressive symptoms and treatment-relevant biological pathways.	Not interoception-specific; broad mechanistic review and heterogeneous clinical evidence.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
23	Pistis et al., 2025	Genetic-factor and cohort association study	CoLaus/PsyCoLaus : 1407 MDD and 2020 controls; NESDA: 1803 MDD and 266 controls	Linked polygenic liability related to energy homeostasis with MDD heterogeneity, particularly atypical symptoms such as increased appetite and hypersomnia.	Predominantly European cohorts; genetic scores explain limited variance; interoception was not directly measured.
24	Mehling et al., 2012	Questionnaire development and psychometric validation	325 adults, predominantly female, White, and highly educated	Developed the eight-domain MAIA to assess subjective bodily attention, appraisal, and regulation.	Convenience sample with restricted demographics and high mind-body experience; self-report only.
25	Eggart et al., 2021	Naturalistic pre-post psychometric study	110 hospitalized MDD patients at baseline; 87 with pre-post data	Supported MAIA-2 applicability and sensitivity to change in hospitalized MDD, although some subscales had subthreshold baseline reliability.	No control group; inpatient sample, attrition, and self-report change limit causal interpretation.
26	Schandry, 1981	Cross-sectional laboratory psychophysiological study	40 adults recruited; 39 analyzed (16 women, 23 men; age 18-40 years)	Better heartbeat perceivers showed higher state anxiety and emotional lability despite similar heart rates; counting itself increased autonomic arousal.	Small nonclinical sample; extreme-group comparison; HCT is vulnerable to timing, prior heart-rate knowledge, and task-induced physiological changes.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
27	Blickle et al., 2024	Cross-sectional multimodal case-control study	30 depressed inpatients and 30 healthy controls aged >50 years	Older depressed inpatients showed no group difference in HCT performance or HRV despite altered insula-related functional connectivity.	Modest older sample; cross-sectional design; medication, cardiovascular, autonomic, and HCT-validity confounding.
28	Tan et al., 2026	Systematic review and meta-analysis	42 studies; HCT analyses included 3295 participants	Depressive symptoms showed a small negative association with HCT performance ($r=-0.10$), but not with heartbeat-discrimination performance; tested moderators were non-significant.	Mostly cross-sectional, heterogeneous samples and task implementations; HCT construct-validity concerns remain.
29	Desmedt et al., 2018	Experimental methodological study	123 healthy university participants	Scores fell by about half when instructions required reporting only genuinely perceived heartbeats, indicating substantial non-interoceptive contribution to standard HCT performance.	Young nonclinical sample; instruction/order effects and subjective strategy reports; no clinical replication.
30	Desmedt et al., 2020	Secondary analysis of two experimental datasets	Two healthy samples (N=194 and N=120)	Time estimation and resting-heart-rate knowledge contributed to HCT performance under standard instructions and were attenuated by adapted instructions.	Non-clinical sample; crude baseline heart rate assessment; adapted instructions risk inducing a conservative response bias.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
31	Zamariola et al., 2018	Methodological correlational analysis	572 participants	Conventional HCT scores were dominated by underreporting, weakly tracked actual heartbeat counts, varied across intervals, and were structurally related to heart rate.	Correlational critique without a perceptual gold standard; does not quantify how much genuine interoceptive signal remains.
32	Zimprich et al., 2020	Commentary and statistical reanalysis	Reanalysis of the Zamariola et al. dataset (N=572)	Reanalysis argued that several apparent HCT problems reflect ratio-score properties and inadequate bivariate analyses rather than complete task invalidity.	No independent sample; statistical reanalysis does not resolve instruction or physiological confounding.
33	Schulz et al., 2021	Experimental cardiovascular-perturbation study	49 healthy adults	Heartbeat-discrimination performance was the most stable predictor of HCT scores during orthostatic change; time-estimation accuracy was not independently predictive.	Small healthy sample; laboratory tilt manipulation; findings are task- and scoring-specific.
34	Shoji et al., 2018	Cross-sectional translation and psychometric validation	390 Japanese young adults	Supported the multidimensional structure and reliability of the Japanese MAIA while identifying population-specific psychometric differences.	Young nonclinical sample; self-report only; some subscale/factor differences limit direct cross-cultural comparison.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
35	Cabrera et al., 2018	Multisample psychometric validation	Three Spanish/US internet and college samples; combined N=1320	Supported one Body Awareness factor and separate supra- and subdiaphragmatic autonomic-reactivity factors for the BPQ-SF.	Convenience/online samples and self-report outcomes; no objective interoceptive criterion.
36	Brand et al., 2024	Multisite psychometric validation	Three German samples; total N=1292	Supported a three-factor German BPQ-SF; autonomic-reactivity scales were reliable, whereas Body Awareness reliability was weaker and reflected both attention and perceived accuracy.	Predominantly nonclinical convenience samples; common-method variance and limited criterion validation.
37	Chen et al., 2025	Cross-cultural adaptation and psychometric validation	Two Chinese coronary-heart-dise ase samples (N=371 and N=350; total N=721)	A culturally adapted BPQ-SF showed acceptable reliability and validity in two Chinese coronary-heart-disease samples.	Medical rather than psychiatric population; culturally specific item revision; self-report and single-country sampling.
38	Murphy et al., 2018	Three cross-sectional experimental studies	Healthy samples with analyzable N=44, 52, and 36	Across three studies, alexithymia was associated with poorer interoception in the respiratory, muscular-effort, and taste domains, independent of autism, depression, and anxiety symptoms.	Small nonclinical samples; novel tasks with limited external validation; cross-sectional design.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
39	Murphy et al., 2017	Developmental narrative review	N/A	Proposes a developmental account linking maturation of interoceptive systems with vulnerability to psychopathology.	Theoretical synthesis; sparse longitudinal developmental evidence and limited MDD-specific data.
40	Trevisan et al., 2019	Systematic review and meta-analysis	66 independent samples; N=7146	Meta-analysis found a small negative association between alexithymia and interoception, with effects differing by interoceptive domain and measure.	High construct and measurement heterogeneity; mostly cross-sectional studies and possible reporting bias.
41	Zhou et al., 2026	Multicohort EEG-ECG case-control study with stress substudy	Three resting datasets totaling N=693; acute-stress substudy N=140	MDD shows a reproducible central-peripheral dissociation: stable central cardiac interoception deficits (reduced HEP) versus reactive peripheral autonomic dysfunctions (elevated HR, reduced HRV) under stress, with central and peripheral measures linked to distinct symptom dimensions. These findings challenge a unitary hierarchical model and support a brain-body framework for MDD.	Cross-sectional cohorts; HEP and autonomic measures remain preprocessing/physiology sensitive; prognostic utility is untested.
42	Wang et al., 2026	Narrative/methodological review	80 studies	Reviews HEP paradigms, analytical choices, clinical findings, and their potential as neural indices of cardiac afferent processing.	Heterogeneous disorders and protocols; no formal biomarker validation; HEPs do not directly measure conscious IAC.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
43	Park et al., 2025	Cross-sectional fMRS and computational-learning study	52 healthy young adults with usable data	Higher anterior-insula Glx predicted greater error sensitivity, which statistically linked Glx with transdiagnostic internalizing/general psychopathology symptoms.	Small healthy sample; single-voxel fMRS and cross-sectional mediation; not MDD-specific.
44	Torous et al., 2021	State-of-the-field narrative review	N/A	Reviews digital phenotyping and remote mental-health assessment/intervention using smartphones, apps, social media, AI, and virtual reality.	Broad digital-psychiatry review; relevance to interoception is indirect and technology evidence is heterogeneous.
45	Sumińska and Rynkiewicz, 2025	Three-arm mixed-design randomized trial	72 randomized; 63 completed (23 smartwatch-MBSR, 20 MBSR, 20 control)	Both MBSR groups reduced stress; smartwatch-supported physiological feedback increased mindfulness but did not significantly improve anxiety or depression.	Nonclinical/high-stress sample; modest groups and attrition; physiological signals were not validated as interoceptive measures.
46	Zhou et al., 2022	Cross-sectional EEG case-control study	27 MDD patients and 27 healthy controls	MDD was associated with poorer heartbeat-counting performance and attenuated HEPs; prestimulus HEP activity predicted subsequent neural responses to sad faces.	Small single-center case-control sample; HCT validity and HEP analytical/cardiac confounding.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
47	Wu et al., 2025	Cross-sectional sex-stratified mediation study	390 patients with MDD (150 men and 240 women)	MAIA dimensions partially mediated the association between perceived social support and depressive symptoms, with different mediating patterns in men and women.	Cross-sectional self-report mediation; single-center MDD sample; no objective interoceptive measure or healthy comparison.
48	Wang et al., 2024	Cross-sectional case-control and mediation study	292 MDD patients and 257 healthy controls	Within MDD, Body Listening, Emotional Awareness, Self-Regulation, and Trusting showed indirect effects linking social support with sociability.	Concurrent self-report mediation cannot establish temporal direction; common-method variance and no objective interoception.
49	Liu et al., 2025	Cross-sectional structural-equation-modeling study	308 patients with MDD	Structural-equation modeling positioned interoception as a central node linking environmental, cognitive, and behavioral-emotional factors in MDD.	Cross-sectional, predominantly self-report SEM; model fit does not establish causality or external prediction.
50	Lv et al., 2025	Cross-sectional case-control study	71 young patients with bipolar disorder and 111 controls	Young patients with bipolar disorder reported lower interoceptive sensibility and greater somatic/emotional symptoms; Not-Distracting remained independently associated with symptom burden.	Bipolar rather than MDD sample; cross-sectional self-report design; medication and maintenance-state effects.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
51	Yu et al., 2025	Multicenter cross-sectional survey	406 patients with chronic heart failure from six county hospitals	Interoceptive self-report scores in chronic heart failure were associated with depressive symptoms, social support, illness perception, hospitalizations, BMI, and symptom burden.	Medical rather than MDD population; convenience sampling and cross-sectional self-report data.
52	Su et al., 2025	Cross-sectional clinical study	73 patients with COPD	In COPD, interoceptive abnormalities and anxiety sensitivity were associated with anxiety symptoms and indices of acute exacerbation.	Small, older, predominantly male medical sample; cross-sectional self-report design; indirect relevance to primary MDD.
53	Mavroudis et al., 2025	Comprehensive narrative synthesis	117 studies across functional neurological disorder subtypes	Narratively synthesized high and variable rates of depression, anxiety, PTSD, dissociation, and medical comorbidity across FND subtypes.	Heterogeneous definitions and samples; no meta-analysis; interoception was not consistently measured.
54	Gninenko et al., 2025	Cross-sectional diffusion-MRI case-control study	85 patients with functional neurological disorder and 75 controls	FND was associated with widespread white-matter microstructural reductions; some group effects attenuated after adjustment for depression/anxiety.	Cross-sectional mixed-FND sample; mood-symptom confounding; indirect relevance to MDD and interoception.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
55	Harris, 2025	Hypothesis and theory article	N/A	Introduces Active Intersubjective Inference, linking recursive social self-models with interoceptive predictions and applying the account to MDD and PTSD.	Conceptual framework; proposed neurocomputational mappings and clinical predictions require empirical testing.
56	Lazaridou et al., 2025	Prospective multicenter observational-study protocol	Planned N=300 coronary-bypass patients (200 development, 100 validation)	Protocol for predicting postoperative depressive symptoms by integrating psychological, autonomic, endocrine, inflammatory, and interoceptive variables.	Protocol only; cardiac-surgery population rather than primary MDD; outcome performance remains unknown.
57	Datko et al., 2022	Single-arm pre-post fMRI mindfulness study	41 recruited; 28 completed usable pre-post imaging	Eight weeks of mindfulness training increased left anterior-insula responses during interoceptive attention; greater neural change was associated with increased Body Trusting.	Single-arm imaging analysis with small completer sample and attrition; mixed anxiety/depression population.
58	Schuman-Olivier et al., 2024	Secondary mediation analysis of a randomized trial	274 primary-care adults randomized; 49% had a depressive disorder	Changes in interoceptive appreciation partially mediated mindfulness-related behavior change; Body Listening and Trusting effects varied with baseline depression severity.	Secondary self-report mediation in mixed diagnoses; behavioral outcome rather than depression remission; mechanism is not proven.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
59	Remiszewski et al., 2025	Randomized controlled exercise trial	62 physically inactive young adults (32 exercise, 30 passive control)	Twelve weeks of aerobic exercise improved HCT-derived accuracy and confidence and reduced depressive/anxiety symptoms, with most gains by 6 weeks.	Nonclinical young sample; passive control; HCT-based interoceptive outcome and self-report symptoms.
60	Silva et al., 2026	Three-condition pilot experimental study	24 healthy adults	Cardiac-music synchronization improved cardiac interoceptive performance; HRV increased and heart rate/negative affect decreased across experimental conditions.	Very small healthy sample; single-session effects and uncertain generalizability to MDD.
61	Cerritelli et al., 2020	Randomized sham-controlled fMRI trial	32 adults with chronic low back pain randomized to OMT or sham	OMT reduced heartbeat-tracking errors and produced later reductions in BOLD responses in bilateral insula, ACC, striatum, and right middle frontal gyrus.	Small chronic-pain sample, not MDD; heartbeat-counting-type task; short follow-up and exploratory mechanistic interpretation.
62	Levy and Jung, 2026	Randomized open-label pilot trial	29 volunteers analyzed (15 intervention, 14 control)	Four weekly myofascial/lymphatic OMT sessions improved body awareness and body dissociation scores and reduced anxiety, depressive symptoms, and stress.	Small open-label volunteer sample from one osteopathic college; no sham control; self-report outcomes and short follow-up.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
63	Kapadia et al., 2019	Uncontrolled pilot feasibility study	10 individuals with moderate-to-severe MDD	Facial functional electrical stimulation was feasible and well tolerated in MDD and was accompanied by reductions in clinician-rated and self-reported depressive symptoms.	Very small uncontrolled pilot; no sham comparison and no interoceptive outcome; strong expectancy/regression-to-mean risk.
64	Chiesa et al., 2025	Nonrandomized controlled pre-post study	38 mindfulness-teacher trainees and 24 matched controls (N=62)	Nine months of mindfulness-teacher training improved selected MAIA dimensions; greater Self-Regulation gains correlated with lower depressive/emotional-distress scores.	Self-selected, nonrandomized professional sample; self-report outcomes; no clinical MDD sample or follow-up.
65	Fridgeirsson et al., 2025	Longitudinal resting-state study	DBS fMRI 9 treatment-resistant MDD patients and 12 healthy controls with usable imaging	One year of ventral anterior-limb internal-capsule DBS altered amygdala-insula/prefrontal and nucleus-accumbens connectivity in treatment-resistant MDD.	Nine-patient highly selected cohort; no sham-controlled long-term clinical comparison; modeling and medication confounding.

Ref.	Study	Study design / evidence type	Population and sample size	Key findings or contribution	Principal limitations
66	Xu et al., 2025	Six-month longitudinal DBS EEG study	8 patients with treatment-resistant MDD	Subcallosal-cingulate DBS increased HEP amplitude, showed acute stimulation on/off modulation, and larger longitudinal HEP increases were associated with faster response.	Eight-patient open-label cohort; no sham group; HEP/stimulation-artifact sensitivity; association does not establish mediation.
67	Bohlen et al., 2025	Single-blind randomized sham-controlled trial protocol	Planned N=60 adults with chronic low-back pain and depressive symptoms	Protocol will test OMT effects on depressive symptoms, pain/function, and interoceptive accuracy, sensibility, and awareness in chronic low back pain.	Protocol only; no outcome data; comorbid pain population rather than primary MDD and blinding is challenging.
68	Bornemann et al., 2014	Psychometric validation and longitudinal contemplative-training study	Validation N=1076; intervention N=232 (152 training, 80 retest controls)	Three months of contemplative training produced domain-specific changes in self-reported interoceptive awareness, especially regulatory aspects of the MAIA.	Healthy self-selected sample; self-report outcomes and bundled training; objective interoception was not assessed.
69	Harrison et al., 2025	Pooled multicenter cross-sectional experimental study	175 participants from four sites (51% women)	State anxiety is linked to reduced metacognitive bias overall; reduced interoceptive insight is associated with anxiety only in women.	Cross-sectional pooled nonclinical data; inspiratory-load task only; limited relevance to depression and binary gender categorization.

References are numbered as in the main manuscript. Sample size is reported for empirical studies where available. For systematic reviews and meta-analyses, the

number of included studies and pooled sample size are reported when provided by the original publication. For nonempirical publications, sample size is reported as not applicable; for study protocols, planned sample sizes are shown. The principal limitations represent a concise critical appraisal relevant to the interpretation of the present review and do not constitute a formal risk-of-bias assessment. ACC: Anterior cingulate cortex; BPQ-SF: Body Perception Questionnaire-Short Form; DBS: Deep brain stimulation; DMN: Default mode network; EEG: Electroencephalography; FES: Functional electrical stimulation; fMRI: Functional magnetic resonance imaging; fMRS: Functional magnetic resonance spectroscopy; Glx: Glutamate-glutamine; GWAS: Genome-wide association study; HCT: Heartbeat counting task; HEP: Heartbeat-evoked potential; HRV: Heart rate variability; IAc: Interoceptive accuracy; ICA: Independent component analysis; MAIA: Multidimensional Assessment of Interoceptive Awareness; MAIA-2: Multidimensional Assessment of Interoceptive Awareness, Version 2; MBSR: Mindfulness-based stress reduction; MDD: Major depressive disorder; MRS: Magnetic resonance spectroscopy; OMT: Osteopathic manipulative treatment; RCT: Randomized controlled trial; SSRI: Selective serotonin reuptake inhibitor; BMI: Body mass index; BOLD: Blood oxygen level-dependent; COPD: Chronic obstructive pulmonary disease; ECG: Electrocardiography; FND: Functional neurological disorder; MRI: Magnetic resonance imaging; NESDA: Netherlands Study of Depression and Anxiety; pgACC: Pregenual anterior cingulate cortex; PTSD: Post-traumatic stress disorder; SANRA: Scale for the Assessment of Narrative Review Articles; SEM: Structural equation modeling; SMN-VN: Somatomotor-visual networks; VAN: Ventral attention network.