Molecular Psychiatry (2020) 25:321–338
https://doi.org/10.1038/s41380-019-0585-z

REVIEW ARTICLE

Prospective biomarkers of major depressive disorder: a systematic review and meta-analysis

Mitzy Kennis[1] · Lotte Gerritsen[1] · Marije van Dalen[1] · Alishia Williams[1,2] · Pim Cuijpers[3] · Claudi Bockting[4,5]

Received: 18 January 2019 / Revised: 9 July 2019 / Accepted: 19 August 2019 / Published online: 19 November 2019

Abstract

Leading biological hypotheses propose that biological changes may underlie major depressive disorder onset and relapse/recurrence. Here, we investigate if there is prospective evidence for biomarkers derived from leading theories. We focus on neuroimaging, gastrointestinal factors, immunology, neurotrophic factors, neurotransmitters, hormones, and oxidative stress. Searches were performed in Pubmed, Embase and PsychInfo for articles published up to 06/2019. References and citations of included articles were screened to identify additional articles. Inclusion criteria were having an MDD diagnosis as outcome, a biomarker as predictor, and prospective design search terms were formulated accordingly. PRISMA guidelines were applied. Meta-analyses were performed using a random effect model when three or more comparable studies were identified, using a random effect model. Our search resulted in 67,464 articles, of which 75 prospective articles were identified on: Neuroimaging ($N = 24$), Gastrointestinal factors ($N = 1$), Immunology ($N = 8$), Neurotrophic ($N = 2$), Neurotransmitters ($N = 1$), Hormones ($N = 39$), Oxidative stress ($N = 1$). Meta-analyses on brain volumes and immunology markers were not significant. Only cortisol ($N = 19$, OR = 1.294, $p = 0.024$) showed a predictive effect on onset/relapse/recurrence of MDD, but not on time until MDD onset/relapse/recurrence. However, this effect disappeared when studies including participants with a baseline clinical diagnosis were removed from the analyses. Other studies were too heterogeneous to compare. Thus, there is a lack of evidence for leading biological theories for onset and maintenance of depression. Only cortisol was identified as potential predictor for MDD, but results are influenced by the disease state. High-quality (prospective) studies on MDD are needed to disentangle the etiology and maintenance of MDD.

Introduction

Major depressive disorder (MDD) is a disabling disorder that is amongst the most prevalent mental health disorders worldwide [1, 2] and is highly recurrent [3–5]. Therapeutic strategies, such as antidepressant medication, are available, although outcomes are suboptimal given roughly 50% of patients do not adequately respond [6, 7]. In order to improve treatment approaches and prevent recurrence, it is important to examine the underlying vulnerabilities that predispose individuals to depression onset and recurrence. By prospectively investigating biological predictors of MDD onset, relapse and recurrence, more insights into the potential causes of MDD can be gained. For these purposes, biomarkers could be particularly informative for understanding the etiology of MDD, and could stimulate development of new clinical approaches in the future.

Numerous studies suggest that MDD is related to alterations in various biological systems [8, 9]. For instance, MDD has been associated with alterations in brain structure

Supplementary information The online version of this article (https://doi.org/10.1038/s41380-019-0585-z) contains supplementary material, which is available to authorized users.

✉ Claudi Bockting
c.l.bockting@amsterdamumc.nl

[1] Department of Clinical Psychology, Utrecht University, Utrecht, The Netherlands

[2] School of Psychology, Faculty of Science, the University of New South Wales, Sydney, NSW, Australia

[3] Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health research institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[4] Department of Psychiatry, Amsterdam University Medical Centers, location AMC, University of Amsterdam, Amsterdam, The Netherlands

[5] Institute for Advanced Study, University of Amsterdam, Amsterdam, The Netherlands

and function, (e.g. [10, 11]), gastrointestinal factors (e.g. [12, 13]), immunology (e.g. [14]), endocrinology (including neurotransmitters, e.g. [15, 16]), neurotrophic factors (e.g. [17, 18]), hormones (e.g. [19]), and oxidative stress (e.g. [20]). Based on these frequently reported biomarker alterations several biological hypotheses for the etiology of MDD have been formulated. Support for these hypotheses have primarily been derived from cross-sectional studies. However, cross-sectional studies cannot provide evidence for *causality*, and thus cannot distinguish causes from consequences secondarily to the illness [21]. To determine whether an etiological mechanism is potentially causal for the development of MDD, the minimal requirement for a study is that the biomarkers are assessed before the development of MDD or prior to a recurrent episode. Thus, prospective studies investigating biomarkers before the onset or relapse/recurrence of MDD are necessary. Further, there are indications that first onset versus relapse/recurrence of MDD is based on different mechanisms [22, 23]. Therefore, investigating predictive biomarkers for onset and relapse/recurrence separately can improve predictive models. However, to our knowledge, no systematic overview of prospective studies comparing biomarkers of onset and relapse/recurrence of MDD has been conducted.

Therefore, we will provide a systematic overview of prospective studies investigating leading biological hypotheses on the etiology of MDD. The first goal is to determine whether there is prospective evidence that these biomarkers predict onset, and relapse/recurrence of MDD. A systematic search for prospective studies will be performed. We explicitly focus on studies using a clinical interview to determine the onset and re-occurrence of a major depressive episode. The search is subdivided into the following biological areas: neuroimaging, gastrointestinal, immunology, neurotrophic, neurotransmitters, hormones, and oxidative stress (see Supplementary Fig. 1). The second goal will be to establish the robustness of each biomarker and to compare the effect size of different biomarkers. Further, subgroup analyses and meta-regression will be performed to investigate potential moderators.

Methods

Search process and study selection

The study was performed according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (PRISMA [24]; see Appendices A and B for search terms and flow charts and Appendix C for PRISMA checklist). This meta-analysis was part of a larger project on evidence for leading theories for MDD onset, and relapse/recurrence and mechanisms of change (for the current study see registration in Prospero CRD42017072990; for psychological predictors of depression see Prospero CRD42017073975; CRD42017073977). Literature searches per biological system were performed between July 2016 and July 2017 in the online databases PubMed, PsychINFO and EMBASE, and a combined search update was performed in June 2019. No start date was included, so all articles that were digitalized until June 2019 were included. The search included terms related to: (1) MDD, (2) longitudinal studies predicting onset, relapse and recurrence, and (3) biological systems of interest (see Appendix A). The articles were independently screened for eligibility based on title and abstract (see criteria below) by two team members, including at least one of the researchers (MK, LG, or MvD), and a member of our screening team (psychology/research Master students; see "Acknowledgements").

The following inclusion criteria were applied: (1) Diagnostic status of MDD for all participants through clinical interview (e.g., SCID, K-SADS from DSM, CIDI from ICD) or report of a clinician-assessed diagnosis (e.g., being hospitalized for MDD treatment, self-report of being diagnosed with MDD by a clinician). (2) The study design is longitudinal. (3) The target variable(s) (e.g., the proposed vulnerability factors) are assessed prospectively, that is before the onset or relapse/recurrence of MDD. (4) The target variable is derived from one of the leading biological models. Exclusion criteria were: diagnosis of mood disorders other than MDD (e.g. bipolar disorder), late-life depression, MDD due to the other (medical) disorders, or studies including a mixed group of diagnoses where less than 75% was diagnosed with MDD. In order to trace studies published after the initial search date, and to add recently published studies, we screened of the included articles the reference list, articles citing, and reference lists of recent reviews. This was done between August and September 2017, and in June 2019 for the new inclusions.

Data extraction and quality assessment

Data extraction was performed by two team members independently (but not blind to the data extracted by the first data extractor) including at least one author (MK, LG, and MvD) and a member of our screening team (see "Acknowledgement"). The following data were extracted: number of included participants and group membership (developing MDD or not), age, gender, study country, MDD diagnosis at baseline, assessment tool of diagnosis, diagnostic criteria, biomarker measurement outcome, biomarker type of measurement, biomarker time of measurement, follow-up time, summary of main outcome. The quality of included studies was assessed by two team members according to a minimally adjusted version of the GRADE guidelines on study level [25]. Information was

extracted on selection of cohorts (similar for groups compared), quality of MDD assessment instrument, presence of baseline MDD (symptoms), matching of samples or adjustion for covariates, biomarker assessment, interviewer, description of drop-outs, description of interventions, and other sources of bias. A score for the quality was also given, by counting the number of questions where there was limited risk of bias (max score = 9).

Analysis

Random effects meta-analyses were performed using comprehensive meta-analysis (www.meta-analysis.com). A meta-analysis was conducted when three or more studies were included using a similar modality of biomarker assessment [26]. When multiple studies investigated the same sample, analysis included only the study with the largest sample size. Odds ratio or risk ratio were the summary effects of outcome. Significance was determined with $p = 0.05$ for meta-analyses. First, analysis was performed on onset and relapse/recurrence of MDD combined to investigate the predictive effect of all biomarkers on MDD development in general. Differences between biomarker effects was also investigated with a subgroup analysis. If a difference exists, meta-analyses were performed per biomarker. Second, separate analyses were performed on studies including participants without baseline clinical MDD diagnosis and/or first onset only versus studies including participants with baseline clinical MDD diagnosis and/or relapse/recurrence (including mixed groups with onset and relapse/recurrence). Heterogeneity was assessed with the Q-test and I^2 statistic [27]. Sensitivity analyses were also performed by re-running analyses after removal of outliers (defined by having no overlap of the 95% CI with the pooled effect 95% CI) and studies with low risk of bias. Baseline age, percentage female participants, biomarker assessment, follow-up time, and quality assessment score were assessed as moderators, when sufficient studies (three per subcategory) were included in the analysis. For analysis of biomarker assessment all effect sizes reported were taken into account. Publication bias was also assessed using Egger's test for asymmetry [28] of the funnel plot and Duval and Tweedie's trim and fill procedure [29].

Results

Search results and quality assessment

The PRISMA flow chart provides an overview of the number of articles screened, included and excluded for all biomarkers combined (see Fig. 1; flow charts per biological system can be found in Appendix B). In total, 67,464 articles were screened for eligibility across all biomarkers.[1] After initial screening, eligibility of 707 articles was assessed based on the full text. In total, only 75 unique prospective studies were identified (see Table 1; [30–104]). Overall, 75 prospective articles were identified on: Neuroimaging ($N = 24$), Gastrointestinal factors ($N = 1$), Immunology/inflammation ($N = 8$), Neurotrophic ($N = 2$), Neurotransmitters ($N = 1$), Hormones ($N = 39$), and Oxidative stress ($N = 1$). In total 39,028,432 participants (median 85, range [9–9275]) were included (Table 1), of which 3267 developed MDD over the follow-up period (median 22, range [3–608]). The median age of study participants was 39 [range 9–66] and the the median percentage of females included was 64% [29–100%]. Follow-up time ranged from 4 months to 22 years, which is adequate for detecting onset, relapse or recurrence (median 3 years). The SCID ($N = 23$) and versions of the (K)SADS ($N = 19$) were the most frequently administered clinical interviews to assess MDD using DSM criteria (DSM $N = 54$) over ICD criteria. Studies describing a clinical diagnosis made by two independent psychiatrist or self-report of hospitalization or diagnosis for MDD were also included incidentally ($N = 7$). Most studies were performed in Western countries (e.g. USA, UK, and Germany, see Table 1). Only 38 studies were identified that excluded participants with baseline clinical MDD diagnosis. First onset of MDD was investigated in 31 studies, relapse/recurrence in 35 studies, and 9 studies included mixed onset and relapse/recurrence samples. Overall, the mean quality score of studies was good (average quality score = 6.3, median 6, range (3–9)), 19 studies had a very low risk of bias (>6 quality score), 26 studies had some risk of bias (5–6 quality score), and 8 studies had high risk of bias (4 or lower quality score). Below, meta-analyses will be described and incidental findings will be discussed narratively (see tables in Supplementary material).

Neuroimaging

Out of the 4210 articles screened for neuroimaging, 21 prospective biomarker studies fulfilled eligibility criteria and the update revealed 3 additional articles (total $N =$ 1952, median $N = 83$, MDD development $N = 420$, median $N = 18$, range for age [6–63], % female [29–100], follow-up time [1–10], QA score (4–9)). However, due to overlap in study samples and heterogeneity in methods applied (e.g. tasks, regions of interest), meta-analysis could only be performed on some hippocampus, amygdala and frontal brain area volumes (see Table 1 and Supplementary Fig. 2). No significant odds ratios were observed for volume of the

[1] Note that this number may include duplicates since articles may be screened two times for different classes of biomarkers.

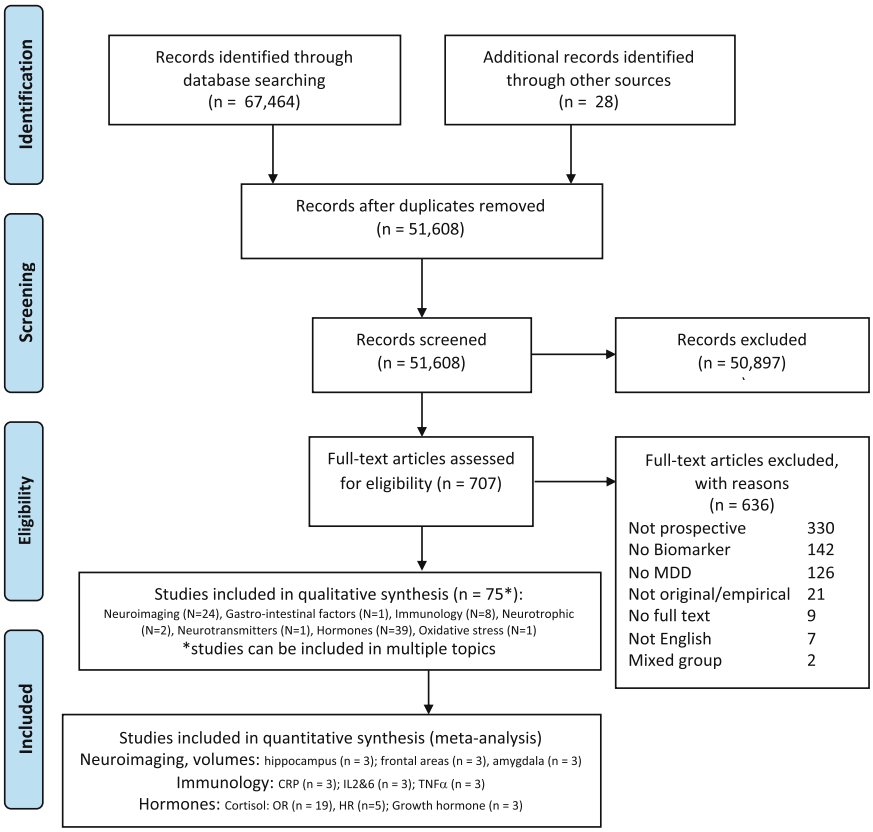

Fig. 1 Flow Diagram of systematic search for prospective studies of MDD overall biological searches combined [24]. See Supplementary material for flow charts per search

hippocampus ($N = 3$, OR = 0.660 [0.426 1.022], $p = 0.063$ [54, 73, 91]), frontal brain regions ($N = 3$, OR = 0.869 [0.480 1.673], $p = 0.730$ [51, 74, 95]), nor the amygdala ($N = 3$, OR = 6.108 [0.143 261.388], $p = 0.345$ [54, 74, 81]). Due to the small number of studies, no further analyses were performed.

Incidental structural MRI studies reported that both smaller and larger insula volume was significantly related to MDD development in two studies [51, 91]. No significant predictive value of the amygdala volume was found in three studies investigating two unique samples [53, 54, 74]. Two studies investigated cortical thickness in the same sample. MDD was predicted by a thinner right para-hippocampus and right fusiform gyrus but not by subcortical thickness [85, 86]. One study reported that higher ACC gray matter volume predicting MDD onset but did not report enough data for analysis [77].

Ten studies investigated if baseline brain activation predicted MDD onset, of which seven used fMRI [49, 50, 76, 82, 99–101] and three used EEG [31, 39, 83]. Studies were too heterogeneous to compare. These studies showed that MDD development was predicted by: lower activity in the frontal lobe in various contexts ([39] reward task loss-gain contrast [83]; rest [71, 82]; go/nogo task, errors; [31] pre- vs posttryptophan depletion), higher activity in the insula ([99] sentence completion increasing in difficulty), higher subgenual anterior cingulate cortex (ACC) temporal and striatal connectivity ([76] self-blame vs other-blame situations) and higher mPFC activity ([50] viewing sad vs neutral movie clips). One study reported no group differences during rest [49]. Differences in subgenual ACC and MFG connectivity were also found in various regions of these networks during rest [71, 101].

Immunology

Out of the 5603 articles screened for immunology, seven met inclusion criteria [43, 46, 61, 69, 87, 88, 94], and one additional study was identified in the update (total $N =$ 27,009, median $N = 2514$, MDD development $N = 1682$, median $N = 160$, range for age (9–66), % female (43–100), follow-up time (3–12), QA score (4–9)). These studies investigated several markers for immunology: C-Reactive Protein, Interleukin-6 (IL-6), IL-1ß, Tumor Necrosis Factor-α (TNFα), Soluble Urokinase Plasminogen Activator Receptor (suPAR), 3-nitrotyrosine, and heat-shock protein 70 (HSP70) in blood or serum

CRP was the investigated in five studies with compatible measures for odds ratio [43, 46, 55, 69, 88, 94], IL (1 and or 6) in four studies, of which two studies investigated the same sample. No significant predictive effects for CRP ($N = 4$, OR = 1.557, 95% CI [0.870 2.788], $p = 0.136$) IL

Table 1 Study details on all prospective studies ($N = 75$), subdivided by the following biological sublevels: Neuroimaging, Gastrointestinal factors, Immunology, Neurotransmitters, Neurotrophic factors, Hormones: HPA axis, HPG axis, HPS axis, and HPT axis

Neuroimaging	Baseline MDD excl. Y = yes, N = No	Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear	Country and cohort information (nr indicates similar cohorts)	Total *N*	Onset (O) or relapse/ recurrence (R) of MDD N	Baseline age (mean or range)	% female	Follow-up time (years; mean or max)	MDD diagnostic interview; Diagnostic criteria	Biomarker of interest	Measure	Technical details	Direction result of biomarker predicting onset, relapse/ recurrence of MDD	Quality score
Bress et al. [39][a]	Y	b	USA	68	16 O	[15–17]	100	2	DISC; ICD-10	Frontal brain areas (ERP)	Activity: EEG	During reward task. comparing loss-gain	↓ Frontal ERP	7
Davey et al. [49][a]	Y	?	Australia	56	8 O	17	45	2	K-SADS-PL; DSM-IV (nm)	Amygdala-sgACC connectivity	RSFC fMRI	Resting state	↑ Amygdala–sACC connectivity	4
Foland-Ross et al. [51][b]	Y	b	USA	33	18 O	13	100	5	K-SADS-PL and SCID; DSM-IV (nm)	32 brain regions were used for support vector machine classification	Structure: MRI	Volumes	The most important classifiers for MDD were ↓ mOFC, PCG, ACC, and ↑ insula	7
Little et al. [73][c]	Y	?	Australia1	99	26 O	13	29	6	K-SADS-PL. MINI; DSM-IV (nm)	Hippocampus, amygdala, OFC and ACC	Structure: MRI	Volumes	↓ Hippocampus NS amygdala, OFC and ACC	6
Little et al. [74]	?	?	Australia1	137	36 O	13	52	6	K-SADS-PL; DSM-IV (nm)	Hippocampus	Structure: MRI	Volumes	↓ hippocampus	5
Nusslock et al. [83][a]	?	b	USA	40	13 O	20	43	3	SADS-C; DSM-IV	Left frontal brain areas (Alpha power)	Activity: EEG	Rest. eyes-open and eyes-closed	↓ Left frontal activity	6
Papmeyer et al. [85][a]	Y	b	UK1	204	19 O	21	56	2	Clinical interview; DSM-IV (nm)	Frontal brain regions	Structure: MRI	Gray matter thickness	↓ Right parahippocampal and fusiform gyrus	9
Papmeyer et al. [86][a]	Y	b	UK1	204	19 O	21	56	2	Clinical interview; DSM-IV	Subcortical brain regions	Structure: MRI	Gray matter thickness	NS	9
Whalley et al. [99][a]	Y	b	UK1	156	20 O	21	52	2	SCID ; DSM-IV	Brain regions activated by the task	Activity: fMRI	Cognitive sentence completion task: increase in difficulty	↑ Insula activity	7
Whalley et al. [100][a]	Y	b	UK1	50	11 O	23	Unclear	2	SCID; DSM-IV	Amygdala, insula, hippocampus, ACC, thalamus	Activity: fMRI	View emotional images vs neutral	↑ Thalamus, insula, ACC activity	7
Nickson et al. [81][c]	N	b	UK1	131	30 O	21	69	4.9	SCID; DSM-IV	VBM	Structure: MRI	Whole brain VBM	↑Amygdala gray msatter volume	7
Macoveanu et al. [77][b]	Y	b	Denmark	85	12 O	39	65	7	SCAN 2.1; (ICD-10)	VBM	Structure: MRI	Whole brain VBM	↑ACC gray matter volume	5
Belden et al. [36][a]	Y	?	USA	129	24 O + R	[6–12]	48	10	Clinician and TRD; DSM-5	Anterior insula	Structure: MRI	Volumes	↓ Insula volume	6
Rao et al. [91][a]	Y	a	USA	83	29 O + R	15	58	5	LIFE. PSR; DSM-IV (at baseline)	Hippocampus	Structure: MRI	Volumes	↓ Hippocampus	5
Serra-Blasco et al. [95][a]	N	a	Spain	24	10 O + R	48	75	5	Life-Chart Manual for Recurrent Affective Illness; Unclear	Whole brain (Freesurfer)	Structure: MRI	Volumes	↓rIFG, ACC, rMFG volumes predict recurrence	8
Allen et al. [31][a]	Y	b	USA	9	3 R	49	61	0.5	SCID; DSM-III-R	Frontal asymmetry	Activity: EEG	Rest. pre- post-tryptophan depletion	↑ Right frontal activity after vs before TD predict lower change of recurrence at follow-up	6
Farb et al. [50][a]	Y	b	Canada	16	10 R	39	69	1.5	SCID; DSM-IV	ROI mPFC	Activity: fMRI	Viewing sad vs natural movie clips	↑ mPFC activity	7
Frodl et al. [53][c]	N	a	Germany1	30	12 R	48	60	1.5	Clinical interview, 2 psychiatrists; DSM-IV	Hippocampus and amygdala	Structure: MRI	Volumes	↓ Hippocampus NS amygdala	6
Frodl et al. [54]	N	a	Germany1	30	13 R	45	63	3	SCID; DSM-IV	Hippocampus and amygdala	Structure: MRI	Volumes	NS	6
Kronmüller et al. [70][b]	N	a	Germany	57	21 R	44	58	2	SCID / LIFE; DSM-IV	Hippocampus	Structure: MRI	Volumes	↓ Hippocampus	5
Lythe et al. [76][a]	Y	b	UK	95	25 R	34	64	1.2	LIFE; DSM-IV	ROI based on previous studies: ACC, temporal, striatal	Activity: fMRI	Activity during self-blame versus other-blame emotions	↑ sgACC and temporal regions connectivity and putamen and claustrum connectivity in recurrent	9

Table 1 (continued)

Neuroimaging	Baseline MDD excl. Y = yes, N = No	Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear	Country and cohort information (nr indicates similar cohorts)	Total *N*	Onset (O) or relapse/ recurrence (R) of MDD N	Baseline age (mean or range)	% female	Follow-up time (years; mean or max)	MDD diagnostic interview; Diagnostic criteria	Biomarker of interest	Measure	Technical details	Direction result of biomarker predicting onset, relapse/ recurrence of MDD	Quality score
Nixon et al. [117][a]	Y	b	UK	38	7 R	[24–63]	33	1	SCID; DSM-IV	ROI based on previous studies	Activity: fMRI	Go/Nogo task	↓ Right dmPFC activity following errors and negative feedback compared to correct hits in recurrence vs other groups	6
Workman et al. [101][a]	Y	?	UK	85	17 R	37	64	1.2	SCID; DSM-IV	sgACC	RSFC: fMRI	Rest. Left sgACC to right sgACC connectivity	Recurrent group was intermediate to resilient and control in sgACC connectivity	9
Langenecker et al. [71][a]	Y	b	USA	94	21 R	21	63	4–6	DIGS and longitudinal interval follow evaluation	fMRI	Activity: fMRI & RSFC fMRI	Go/No-Go task: successful vs unsuccessful inhibition	Successful regulation sgACC hyperactivation. Failed regulation; MFG hypoactivation. RSFC: altered MFG and sgACC connectivity.	8
Gastrointestinal factors	**Baseline MDD excl. Y = yes, N = No**	**Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear**	**Country and cohort information (nr indicates similar cohorts)**	**Total *N***	**Onset (O) or relapse/ recurrence (R) of MDD (N)**	**Baseline age (mean or range)**	**% female**	**Follow-up time (years; mean or max)**	**MDD diagnostic interview; Diagnostic criteria**	**Biomarker of interest**	**Measure**	**Technical details**	**Direction result of biomarker predicting onset, relapse/ recurrence of MDD**	**QA score**
Campo et al. [40] °	Y	?	USA	119	14 O	[6–12]	?	7.5	K-SADS-E; DSM-IV (nm)	L-5-hydroxytryptophan injection at baseline	Abdominal discomfort, nausea, or vomiting in response to the L-5HTP infusion; GI distress	L-5HTP infusion; Survival curve (groups high vs. low sensitive)	↑GI distress after serotonin challenge	6
Immunology	**Baseline MDD excl. Y = yes, N = No**	**Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear**	**Country and cohort information (nr indicates similar cohorts)**	**Total *N***	**Onset (O) or relapse/ recurrence (R) of MDD (N)**	**Baseline age (mean or range)**	**% female**	**Follow-up time (years; mean or max)**	**MDD diagnostic interview; Diagnostic criteria**	**Biomarker of interest**	**Measure**	**Technical details**	**Direction result of biomarker predicting onset, relapse/ recurrence of MDD**	**QA score**
Chocano-Bedoya et al. [43][c]	N	?	USA	4403	81 O	56	100	12	Self-report of diagnosis; Unclear	CRP, IL-6, TNFα-R2	Blood, 1 time point	hs-CRP IA. EIA	NS	4
Haastrup et al. [61][a]	N	?	Denmark	9275	22 O	[18–47]	48	5	prescriptions for antidepressant medication or had a hospital discharge diagnose with the codes F.32 or F.33; ICD-10	suPAR	Blood plasma, 1 time point	ELISA	↑Increased risk (shorter time) to onset of depression with higher SuPAR	6
Rudaz et al. [94]	Y	?	Switzerland1	1524	192 O	51	43	5.5	Diagnostic Interview for Genetic Studies (DIGS); DSM-IV	CRP, IL-1ß, IL-6, TNFα	Blood serum, 1 time point	hs-CRP IA. multiplexed particle-based flow cytometric cytokine assay	↓TNFα, NS for CRP, IL-1ß, IL-6	8
Copeland et al. [46][c]	Y	b	USA	5810	169 O + R	14	49	12	CAPA, YAPA; DSM-IV	CRP	Blood spot, 1 time point	hs-CRP IA	NS	9
Glaus et al. [55][c]	Y/N	?	Switzerland1	2580	608 O + R	[35–66]	54	5.5	DIGS	CRP, IL-6, TNFα	Serum, 1 time pont	IA and latex HS	O + R: ↓TNFα	7
Khandaker et al. [69][c]	N	?	UK	2447	422 O + R	9	51	9	CIS-R; ICD-10	CRP, IL-6	Blood serum, 1 time point	hs-CRP essay. ELISA	NS	5
Pasco et al. [87][a]	Y	?	Australia	822	151 O + R	49	100	10	SCID; DSM-IV-TR	CRP	Blood serum, 1 time point	hs-CRP IA	↑ CRP reduced time till depression (HR)	7

Table 1 (continued)

Neuroimaging	Baseline MDD excl. Y = yes, N = No	Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear	Country and cohort information (nr indicates similar cohorts)	Total *N*	Onset (O) or relapse/ recurrence (R) of MDD N	Baseline age (mean or range)	% female	Follow-up time (years; mean or max)	MDD diagnostic interview; Diagnostic criteria	Biomarker of interest	Measure	Technical details	Direction result of biomarker predicting onset, relapse/ recurrence of MDD	Quality score
Neurotransmitters	**Baseline MDD excl. Y = yes, N = No**	**Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear**	**Country and cohort information (nr indicates similar cohorts)**	**Total *N***	**Onset (O) or relapse/ recurrence (R) of MDD (N)**	**Baseline age (mean or range)**	**% female**	**Follow-up time (years; mean or max)**	**MDD diagnostic interview; Diagnostic criteria**	**Biomarker of interest**	**Measure**	**Technical details**	**Direction result of biomarker predicting onset, relapse/ recurrence of MDD**	**QA score**
Johnston et al. [68][a]	N	a	UK	31	20 R	47	71	8	SCID I/P; DSM-III-R	Plasma Norepinephrine, cortisol	Blood plasma, 1 time point	Chromatography, RIA	↓ Norepinephrine, shorter time to first recurrence.	6
Neurotrophic factors and oxidative stress	**Baseline MDD excl. Y = yes, N = No**	**Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear**	**Country and cohort information (nr indicates similar cohorts)**	**Total *N***	**Onset (O) or relapse/ recurrence (R) of MDD (N)**	**Baseline age (mean or range)**	**% female**	**Follow-up time (years; mean or max)**	**MDD diagnostic interview; Diagnostic criteria**	**Biomarker of interest**	**Measure**	**Technical details**	**Direction result of biomarker predicting onset, relapse/ recurrence of MDD**	**QA score**
Pasquali et al. [88][a]	N	b	USA	148	37 O	40	100	3	SCID; DSM-IV	*Neurotrophic*: BDNF *Immunology:* HSP70, 3-Nitrotyrosine,, *Oxidative stress:* Protein carbonyl, Lipid peroxidation, Thiol content	Blood serum, 1 time point	ELISA, quantitative sandwich enzyme immunoassay, colorimetric assay (thiol)	↑ heat-shock protein 70, 3-nitrotyrosine, protein carbonyl, and lipid peroxidation ↓BDNF	6
Vinberg et al. [97][a]	Y	?	Denmark	234	24 O	44	65	7.5	SCAN; ICD 8/ICD-10	BDNF	Blood, 1 time point	Two-site sandwich ELISA	NS	5
Hormones: HPA	**Baseline MDD excl. Y = yes, N = No**	**Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear**	**Country and cohort information (nr indicates similar cohorts)**	**Total *N***	**Onset (O) or relapse/ recurrence (R) of MDD (N)**	**Baseline age (mean or range)**	**% female**	**Follow-up time (years; mean or max)**	**MDD diagnostic interview; Diagnostic criteria**	**Biomarker of interest**	**Measure**	**Technical details**	**Direction result of biomarker predicting onset, relapse/ recurrence of MDD**	**QA score**
Adam et al. [30][c]	N	?	USA	230	18 O	17	75	1	SCID; DSM-IV	AUC, CAR, diurnal, slope	Saliva, 3 days, 6 times a day	DELFIA	NS	7
Colich et al. [45][c]	Y	b	USA1	89	31 O	12	100	6	K-SADS. SCID; DSM-IV	Cortisol pre- post- trier social stress test	Saliva, 4 time points	LIA	↓ cortisol reactivity in early puberty ↑cortisol reactivity in late puberty	9
Goodyer et al. [56]	?	?	UK2	171	30 O	14	59	1	K-SADS; DSM-IV	Peaks 8:00, cortisol and DHEA	Saliva 4 days, 2 time points	ELISA	↑ cortisol & DHEA	6
Goodyer et al. [57]	?	?	UK2	234	31 O	14	53	1	K-SADS; DSM-IV	Morning and evening cortisol, DHEA	Saliva 4 days, 2 time points	ELISA	↑ DHEA, NS cortisol	7
Goodyer et al. [58][c]	?	b	UK2	367	32 O	14	46	1	K-SADS; DSM-IV	Morning cortisol	Saliva 4 days, 1 time point	ELISA	↑ cortisol	6
Goodyer et al. [59]	?	b	UK2	357	40 O	14	47	1	K-SADS; DSM-IV	Morning cortisol	Saliva 4 days, 1 time point	ELISA	↑ cortisol	6
Grynderup et al. [60][c]	Y	?	Denmark	2920	62 O	56	78	2	SCAN; ICD-10-DSR	Morning, diurnal, evening, morning to evening slope in cortisol	Saliva, 2 hours after awakening, and between 5PM and 4AM	RIA	↓ difference in morning to evening cortisol	9
Harris et al. [63][c]	?	?	UK	116	28 O	39	100	1	SCAN; DSM-IV	DHEA and cortisol	Saliva, 4 days, 2 time points	ELISA	NS	6
Herbert et al. [65][c]	Y	?	UK	279	53 O	37	100	1	SCAN; DSM-IV	Morning cortisol	Saliva 4 days, 1 time point	ELISA	↑ cortisol predicts MDD	8

Table 1 (continued)

Neuroimaging	Baseline MDD excl. Y = yes, N = No	Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear	Country and cohort information (nr indicates similar cohorts)	Total *N*	Onset (O) or relapse/ recurrence (R) of MDD N	Baseline age (mean or range)	% female	Follow-up time (years; mean or max)	MDD diagnostic interview; Diagnostic criteria	Biomarker of interest	Measure	Technical details	Direction result of biomarker predicting onset, relapse/ recurrence of MDD	Quality score
LeMoult et al. [72]	Y	b	USA1	62	26 O	12	100	6.5	K-SADS. SCID; DSM-IV	Morning cortisol	Saliva, 2 days, 4 time points	ELISA	↑ cortisol predicts MDD	9
Rao et al. [93][c]	Y	b	USA2	89	14 O	15	58	5	K-SADS; DSM-IV	Saliva, NUFC	Saliva, 3 time points before sleep and Urine 1 time point before sleep	RIA	↑ cortisol predicts MDD	6
Carnegie et al. [41][c]	N	?	UK	841	46 O + R	15	49	3	CIS-R; ICD-10	AUC, CAR, DD, bedtime (m), waking (m), TEC	Urine, over 24 h	EIA	NS	5
Vrshek-Schallhorn et al. [98][c]	Y	?	USA	270	42 O + R	17	72	4	SCID; DSM-IV	AUC, CAR, diurnal slope	Saliva, 3 days, multiple time points	Time-resolved fluorescent-detection IA	↑ cortisol recurrence	9
Appelhof et al. [32][c]	N	a	Spain	45	22 R	50	44	22	At baseline: SCID. Relapse: HRSD. MADRS. and BD; DSM-IV	Post dex cortisol, Max ACTH, Delta ACTH, Max cortisol, delta cortisol, discharge, difference cortisol	Blood, 2 days, multiple time points before and after DEX/CRH combined with TRH	luminescen-ce enzyme IA	↑ maximal cortisol after DEX/ CRH predicts relapse	6
Aubry et al. [34][b]	Y	a	Switzerland	34	12 R	44	56	1	MINI; DSM-III-R/IV. ICD-10	Cortisol after DEX/CRH test	Blood, 2 days, multiple time points before and after DEX/CRH	Immulite 2000 analyser	↑ AUC and delta in relapse	8
Banki et al. [35][b]	N	a	Hungary	24	9 R	51	100	0.5	Hospital diagnosis; DSM-III-R	CRH, SRIF	CSF, 1 time point	Sensitive and specific IA	↑ CRF in relapse	3
Bockting et al. [37][c]	Y	b	Netherlands1	55	43 R	44	73	5.5	SCID; DSM-IV	Morning and evening cortisol	Saliva, 2 days, 1–2 time points a day	RIA	↓ cortisol related to relapse	7
Bouhuys et al. [38][c]	N	b	Netherlands	77	21 R	44	66	2	CIDI; DSM-IV	24 h urine		RIA	NS	6
Charles et al. [42][c]	N	?	Belgium	13	7 R	[33–67]	77	1.5	SADS-L; DSM-III and RDC	Morning after DST	Blood plasma, 2 time points after taking DEX	RIA	↑ cortisol (non suppression at recovery) higher rates of MDD readmission	4
Chopra et al. [44][b]	Y	b	Canada	55	28 R	39	64	1.5	SCID; DSM-IV	Morning/evening (before mood induction)	Saliva, 4 time points	EIA	↑ cortisol	4
Cosgriff et al. [48][c]	N	b	New Zealand	13	4 R	51	54	0.25	Clinical readmission; Unclear	Mean cortisol, Delta TSH	Blood, 10–12 time points	EIA/RIA	↑ cortisol	5
Hardeveld et al. [62][c]	Y	a	Netherlands	549	193 R	45	71	4	CIDI; DSM-IV	Salivary CAR, evening, DST	Saliva, 2 days, 6 times day 1, 1 time day 2 after DEX	EIA	AUC increase differed, and related to time to recurrence. Other measures, mean evening, DST and AUC were not predictive	7
Hatzinger et al. [64][b]	N	a	Switzerland	20	12 R	52	70	1	SCID; ICD-10	Cortisol after DEX/CRH	Blood plasma, 1 time point after DEX/ CRH	RIA	DEX/CRH test	7
Lok et al. [75]	Y	b	Netherlands1	187	102 R	44	68	2	SCID; DSM-IV	Morning and evening cortisol	Saliva, 2 days, 3 time points	RIA	NS	8
Mander et al. [78][c]	N	?	UK	70	32 R	Unclear	Unclear	3	SCID; DSM-III	DEX suppression	Saliva, 1 day, 3 time points day after DEX	RIA	NS	5
Mocking et al. [118][c]	Y	b	Netherlands1	187	154 R	44	68	10	SCID; DSM-IV	Cortisol/ DHEAS ratio	Saliva, 2 days, 2 time points a day.	RIA	↓DHEAS diurnal course, ↑ cortisol/DHEAS ratio diurnal course	6
Morris et al. [80][c]	Y	a/b	USA	32	9 R	23	63	0.75	SCID; DSM-IV	Cortisol pre- post- TSST	Saliva, 6 time points	ELISA	NS	8
Pintor et al. [89]	N	a	Spain1	43	18 R	51	53	2	SCID; DSM-IV	Cortisol and ACTH after CRF injection	Blood plasma, 6 time points around CRF injection	EIA/RIA	NS Cortisol and ACTH AUC, ACTH after CRF	6

Table 1 (continued)

Neuroimaging	Baseline MDD excl. Y = yes, N = No	Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear	Country and cohort information (nr indicates similar cohorts)	Total *N*	Onset (O) or relapse/ recurrence (R) of MDD N	Baseline age (mean or range)	% female	Follow-up time (years; mean or max)	MDD diagnostic interview; Diagnostic criteria	Biomarker of interest	Measure	Technical details	Direction result of biomarker predicting onset, relapse/ recurrence of MDD	Quality score
Pintor et al. [90][c]	N	a	Spain1	43	18 R	51	46	2	SCID; DSM-IV	Cortisol and ACTH after CRF injection	Blood plasma, 6 time points around CRF injection	EIA/RIA	↑ cortisol (NAUCC) after CRF, ↓ACTH after CRF	5
Rao et al. [92]	N	b	USA2	47	20 R	15	64	3.5	K-SADS-PL; DSM-IV	NUFC	Urine before and after sleep	RIA	↑ cortisol	5
Tsuru et al. [96][c]	N	a	Japan	25	9 R	41	64	10	SCID; DSM-IV	ACTH and cortisol, TSH	Blood, 2 days, day 1 TRH test 5 time points, day 2 DEX /CRH-test, 10 time points	IRMA	NS cortisol and ACTH,↑TSH after TRH test in recurrence	5
Zimmerman et al. [102][c]	N	?	USA	165	47 R	40	73	0.5	Unclear; DSM-III	DEX suppression	Blood, 2 time points after DEX	RIA	NS	6
Zobel et al. [103]	N	a	Germany2	40	10 R	50	65	0.5	Unclear; DSM-IV	DEX/CRH test	Blood, 9 time points DEX/CRH-test	Unclear	↑ cortisol after DEX/CRH at discharge predicts MDD relapse	3
Zobel et al. [104][c]	N	a	Germany2	74	13 R	50	59	0.5	Unclear; DSM-IV	DEX/CRH test, cortisol and ACTH	Blood, 5 time points DEX/CRH-test	RIA	↑ cortisol predicts MDD, ACTH NS	5
Hormones: HPG	**Baseline MDD excl. Y = yes, N = No**	**Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear**	**Country and cohort information (nr indicates similar cohorts)**	**Total *N***	**Onset (O) or relapse/ recurrence (R) of MDD (N)**	**Baseline age (mean or range)**	**% female**	**Follow-up time (years; mean or max)**	**MDD diagnostic interview; Diagnostic criteria**	**Biomarker of interest**	**Measure**	**Technical details**	**Direction result of biomarker predicting onset, relapse/ recurrence of MDD**	**QA score**
Asselmann et al. [33][a]			Germany	1760	165 O	45	50	9	DIA-X/M-CIDI	Testosterone, Androstenedione, sex hormone-binding globuline	Blood serum 8AM-7PM	Liquid-chromatography-tandem mass spectrometry and RIA	NS	6
Hormones: HPS axis	**Baseline MDD excl. Y = yes, N = No**	**Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear**	**Country and cohort information (nr indicates similar cohorts)**	**Total *N***	**Onset (O) or relapse/ recurrence (R) of MDD (N)**	**Baseline age (mean or range)**	**% female**	**Follow-up time (years; mean or max)**	**MDD diagnostic interview; Diagnostic criteria**	**Biomarker of interest**	**Measurement**	**Technical details**	**Direction result of biomarker predicting onset, relapse/ recurrence of MDD**	**QA score**
Coplan et al. [47][c]	Y	?	USA	34	13 O	15	52	9.6	K-SADS & K-SADS-E & SADS-LA; RDC	Growth hormone (sleep)	Blood over 2 nights, 72 time points	RIA	↑GH secretion	8
Franz et al. [52][c]	N	a	USA3	29	22 R	40	100	3	SADS; RDC	Growth hormone (sleep)	Blood before onset of and during sleep, every 20 min	RIA	↑GH secretion	5
Jarrett et al. [66][b]	N	a	USA3	29	22 R	40	100	3	SADS; RDC	Growth hormone (sleep)	Blood before onset of and during sleep, every 20 min	RIA	NS	6
Owashi et al. [84][c]	N	b	Japan	26	6 R	57	76	0.5	Unclear; DSM-IV	Growth hormone, ACTH, cortisol	Blood after DEX/ CRH and GHRH test, 5 time points	RIA	↓GH after GHRH at time of discharge. Cortisol and ACTH around DEX/CRH: NS	4
Hormones: HPT axis	**Baseline MDD excl. Y = yes, N = No**	**Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear**	**Country and cohort information (nr indicates**	**Total *N***	**Onset (O) or relapse/ recurrence (R) of MDD (N)**	**Baseline age (mean or range)**	**% female**	**Follow-up time (years; mean or max)**	**MDD diagnostic interview; Diagnostic criteria**	**Biomarker of interest**	**Measurement**	**Technical details**	**Direction result of biomarker predicting onset, relapse/ recurrence of MDD**	**QA score**

Table 1 (continued)

Neuroimaging	Baseline MDD excl. Y = yes, N = No	Baseline symptoms a = above cut-off, b = below cut-off, ? = unclear	Country and cohort information (nr indicates similar cohorts)	Total *N*	Onset (O) or relapse/ recurrence (R) of MDD N	Baseline age (mean or range)	% female	Follow-up time (years; mean or max)	MDD diagnostic interview; Diagnostic criteria	Biomarker of interest	Measure	Technical details	Direction result of biomarker predicting onset, relapse/ recurrence of MDD	Quality score
			similar cohorts)											
Joffe et al. [67][a]	N	a	Canada	75	71 R	38	66	10	SADS-L; RDC	T4, T3, TSH	Unclear	Unclear	↓T3 was significantly related to time to recurrence. T4, TSH NS	4

Studies are sorted by the inclusion of participants with onset (O), relapse and recurrence (R) of MDD or both (O + R). The second and third columns represent information on the certainty of including healthy individuals at baseline, where the second column shows if baseline MDD diagnosis was excluded with a clinical interview at baseline, and the third column represents whether symptoms were measured with questionnaires (e.g. Hamilton depression scale, beck depression inventory) and whether participants scored above or below the questionnaires cut-off for clinical symptoms at baseline. Basic information on the demographics of participants and study measures, technical details and outcome are also given. The quality score (range 0–9) represents the overall quality of the studies, where a score > 6 represents good quality, indicative of a low risk of bias and < 4 represents poor quality, a high risk of bias.

Specific abbreviations will be mentioned per subsection

General abbreviations: *DIGS* Diagnostic Interview for Genetic Studies, *DSM* diagnostic statistical manual of mental disorders, *ICD* International Statistical Classification of Diseases and Related Health Problems, *MINI* mini-international neuropsychiatric interview, *nm* not mentioned, *NS* nonsignificant, *RDC* research diagnostic criteria, *SCAN* schedules for clinical assessment in neuropsychiatry, *SCID* Structured Clinical Interview for DSM, *K-SADS* Kiddie Schedule for Affective Disorders and Schizophrenia.

Neuroimaging abbreviations: *ACC* anterior cingulate cortex, *dmPFC* dorsal medial prefrontal cortex, *EEG* electroencephalography, *f* functional, *IFG* inferior frontal gyrus, *MFG* middle frontal gyrus, *mOFC* medial orbitofrontal cortex, *MRI* magnetic resonance imaging, orbitofrontal cortex (OFC), *PCG* precentral gyrus, *ROI* region of interest, *RSFC* resting state functional connectivity, *sg* subgenual

Gastrointestinal abbreviations: *GI* gastrointestinal, *L-5HTP* L-5-Hydroxytryptophan

Immunology abbreviations: *CAPA* Child and Adolescent Psychiatric Assessment, *CRP* c-reactive protein, *ELISA* enzyme-linked immunosorbent assay, HR hazard ratio, *IL* interleukin, *SRIF* somatostatin, *TNF* tumor necrosis factor, *YAPA* Young Adult Psychiatric Assessment, *IA* immunoassay

Neurotransmitters abbrevations: *RIA* radioimmunoassay

Neurotrophic and oxidative stress abbreviations: *ELISA* enzyme-linked immunosorbent assay, *BDNF* brain-derived neurotrophic factor

HPA abbreviations: *ACTH* adrenocorticotropic hormone, *AUC* area under the curve, *CAR* Cortisol awakening response, *CRH* Corticotropin-releasing hormone, *DEX/CRH* combined dexamethason-cortisol releasing hormone test, *DST* dexamethasone suppression test, *DHE* adehydroepiandosterone, *ELISA* enzyme-linked immunosorbent assay, *HRSD* Hamilton Rating Scale for Depression, *IA* immunoassay, *LIA* line immunoassay, *MADRS* montgomery-Asberg Depression Rating Scale, *NUFC* nocturnal urinary free cortisol, *RIA* radioimmunoassay, *TSST* trier social stress test.

HPG abbreviations: DIA-X/M-CIDI Munich-Composite International Diagnostic Interview, *RIA* radioimmunoassay

HPS: *GH* growth hormone, *GHRH* growth hormone resleasing hormone, *RIA* radioimmunoassay

HPT abbreviations: *T3* triiodothyronine, *T4* thyroxine, *TSH* thyroid stimulating hormone

[a]Not enough studies to meta-analyze

[b]Not enough data reported in this study to include in meta-analysis

[c]Included in meta-analysis

($N = 3$, OR = 1.025, 95% CI [0.782 1.345], $p = 0.856$) was found. Due to the small number of studies, no further analyses were conducted.

Incidental findings were also identified. One study investigated hazard ratio and showed that CRP significantly predicted earlier time to onset or relapse/recurrence of depression [87]. In three studies (of which two investigated the same sample) TNFα was not found to predict non-significant were also reported [43, 55, 94]. A protein marker for inflammation SuPAR was found to predict reduced time to MDD [61]. In addition, three-nitrotyrosine and HSP70 were higher at baseline in participants that develop vs that do not develop MDD [88].

Gastrointestinal biomarkers

Out of the 760 articles screened for the gut-related biomarkers, only one study met our inclusion criteria [40]. The study showed that children reporting symptoms of abdominal discomfort (e.g. nausea or vomiting) in response to tryptophan (L-5HTP) infusion have a higher risk of developing MDD than children who do not report these symptoms.

Hormones

Out of the 17,114 articles screened, 38 articles were included and 1 study was identified with the update. The studies investigated the following hormonal axes: 35 hypothalamic-pituitary axis (HPA axis; the feedback loop regulation stress responses, including ACTH, CRH, CRF, cortisol), 5 hypothalamic-pituitary-gonadal-axis (HPG-axis: regulating the reproductive system e.g. DHEAS), 4 hypothalamic-pituitary-somatic axis (HPS axis: mainly regulating growth and includes growth hormone (GH)), and 3 hypothalamic-pituitary-thymus-axis (HPT axis; mainly regulating metabolism e.g. thyroid hormone). Results will be described below by these biological/hormonal axes.

HPA axis

The predictive value of cortisol on subsequent MDD was investigated in 35 prospective studies (total $N = 7823$, median $N = 74$, MDD development $N = 1236$, median $N = 26$, range for age (12–56), % female (44–100), follow-up time (1–22), QA score (3–9)). Cortisol was primarily measured in saliva, but differed in time of day of measurement (morning, evening, diurnal, nocturnal, reactivity), and both single time point and multiple time point measurements were included. Cortisol was a significant predictor of subsequent MDD with a small effect size ($N = 19$, OR = 1.294, 95% CI [1.035 1.616], $p = 0.024$ [30, 32, 37, 41, 42, 45, 48, 58, 60, 63, 65, 78, 80, 84, 90, 92, 96, 102, 104], see Fig. 2) overall comparible studies on unique samples. Heterogeneity was large and significant (76%, $p < 0.001$). The effect became nonsignificant when outliers were removed (OR = 1.228; $p = 0.052$) or low quality studies were removed (QA < 4; OR = 1.206, $p = 0.094$). Inspection of the funnel plot showed indication of publication bias (7 studies were missing on the left side), though the Eggers test was not significant $p > 0.05$. Correction for publication bias led to a nonsignificant effect. Further, the quality score of the studies moderated the effect ($\beta = -0.176$, $p = 0.012$) indicating a lower study quality is related to an increased effect size.

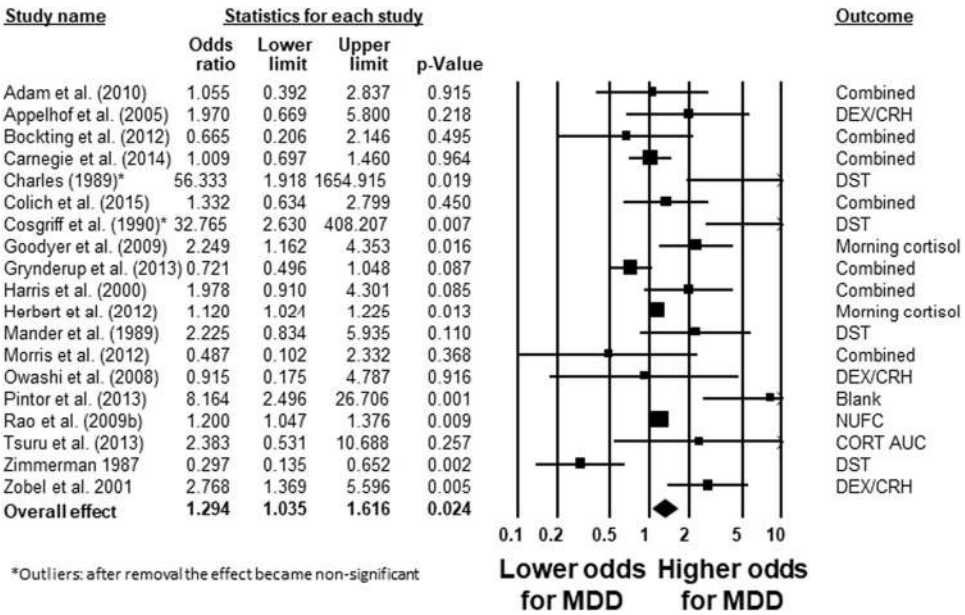

Fig. 2 Forest plot of a meta-analysis on studies investigating measures of cortisol before MDD onset, relapse or recurrence. Charles et al. [42] and Cosgriff et al. [48] are identified as outliers, and excluding them from analysis resulted in a nonsignificant effect

Comparing studies including participants with baseline MDD/mixed group versus no baseline MDD showed a significant higher effect size in the first group ($p = 0.027$), confirming the significance of including baseline clinical MDD diagnosis in studies (disease state effect). The pooled odds ratio for studies including baseline diagnosis was medium and significant ($N = 13$, OR = 1.919, 95% CI [1.072 1.231], $p = 0.012$), while studies excluding baseline diagnosis had a small nonsignificant pooled odds ratio ($N = 6$, OR = 1.082, 95% CI [0.938 1.249], $p = 0.280$). Comparing studies investigating onset, relapse or recurrence, or a mixed groups not significant ($p = 0.107$).

Studies investigating time until MDD onset, relapse or recurrence using Hazard ratios showed no significant predictive effect of cortisol (HR = 1.011, 95% CI [0.963 1.040], $p = 0.447$ [32, 38, 62, 79, 98]). Due to the small number of studies, no further analyses were conducted.

Besides cortisol, other HPA-axis markers in relation to relapse or recurrence of MDD were investigated incidentally. Nonsignificant findings were reported for adrenocorticotrophic hormone (ACTH) [32, 84, 89, 96, 104], and cortisol releasing hormone (CRH; [35]). One study reported lower ACTH in reaction to a DEX/CHR predicts relapse [90]. Thus, it remains unclear if HPA-axis biomarkers predict MDD development or whether results reflect disease state or quality of studies.

HPG axis

HPG biomarkers were investigated in five studies (total $N = 2468$, median $N = 187$, MDD development $N = 408$, median $N = 31$, range for age (14–45), % female [50–100], follow-up time (1–10), QA score (6–7)). Four studies investigated dehydroepiandrosterone (DHEA) or DHEA-sulfate, (DHEAS) in saliva [56, 63, 79], but studies included the same sample and included OR and HR measures, which are not comparable. Both significant predictive effects [56, 57] as well as no significant predictive effects [63] were reported. One study showed that a higher cortisol: DHEAS ratio predicted a shorter time to recurrence [79]. One study investigated serum concentratioins of testosterone, androstenedione, and sex hormone-binding globuline (SHBG) and found no predictive effect on first onset nor the combination of onset/recurrence over 17 years [33]. Thus, it remains unclear if androsterones predict MDD development.

HPS axis

Four studies [47, 52, 66, 84] investigated the predictive effect of GH on subsequent MDD (total $N = 118$, median $N = 29$, MDD development $n = 23$, median $N = 22$, range for age [15–57], % female [52–100], follow-up time (0.5–9.6), QA score (4–8)), of which 2 investigated the same sample and one study that did not provide sufficient data for analysis [47]. Three studies investigated GH secretion over night and a steeper increase in GH secretion was found in participants that had later onset [47] and recurrence [52] of MDD, but another study (on the same sample) found no significant predictive value for recurrence [66], and lower GH is also reported in individuals that relapse [84]. No differences were found in somatostatin (GH releasing factor) in CSF between relapsing and not relapsing participants [35]. Thus, it remains unclear if HPS markers predict MDD development.

HPT axis

Three studies reported results investigating the HPT axis (total $N = 113$, median $N = 25$), MDD development $n = 84$, median $N = 9$, range for age [38–51], % female [54–66], follow-up time [0.25–10], QA score (4–5); [48, 67, 96]. Higher thyroid stimulating hormone (TSH) was related to recurrence in one study [96], but was also found to not differ between people with and without relapse in another study [48]. One study investigated T4, T3, and TSH using cox regression survival analyses, and reported that lower T3 was related to shorter time until relapse/recurrence [67]. Thus, the relation with HPT axis and subsequent MDD remains unclear and study quality was low.

Oxidative stress

Out of the 1336 articles screened, 1 article met inclusion criteria [88]. Pasquali et al. [88] investigated markers for oxidative stress in blood (see Table 1). Lipid peroxidation was higher in participant that develop MDD ($N = 37$) compared to participants who did not develop MDD ($N = 111$). No significant differences between these groups were found for protein carbon and thiol content. Thus, whether oxidative stress predicts subsequent MDD remains unclear.

Discussion

A systematic search for prospective studies investigating biomarkers of MDD onset, relapse, and recurrence was performed. Of the 67,464 articles screened, only 75 prospective studies were identified that studied biomarkers before MDD onset or relapse/recurrenc. Of those, only 38 studies reported results on participants that were healthy (had no MDD diagnosis) at baseline and are thus unconfounded by disease state. Prospective evidence for the majority of biomarkers predicting onset, and relapse/recurrence of MDD was scarce ($N = 75$) and spread over a wide range of topics: Neuroimaging ($N = 24$), Gastrointestinal

factors ($N = 1$), Immunology ($N = 8$), Neurotrophic ($N = 2$), Neurotransmitters ($N = 1$), Hormones ($N = 39$), and Oxidative stress ($N = 1$). Marked heterogeneity across studies was observed for neuroimaging studies ($N = 24$). These included assessments based on EEG, task-based functional MRI, and structural MRI that focused on different brain regions, thereby precluding the calculation of an overall effect [105]. This highlights the urgent need for standardized methods in order to be able to compare data from different samples. The only significant biomarkers that increased odds for MDD onset, and relapse/recurrence was cortisol. However, the inclusion of baseline clinical diagnosis was shown to influence this effect. Therefore, the effect of disease state cannot be ruled out. Meta-analyses on CRP, TNFα, IL2&6, GH, hippocampus, amygdala, and frontal brain areas volume were not significant, potentially due to the limited amount of studies included in these analyses [range 3–4]. Only incidental (<3) studies investigated TSH, DHEAS, amygdala volumes, neurotrophic factors, oxidative stress, ACTH, neurotransmitters and gastrointestinal biomarkers. In addition, results on biomarkers were inconsistent.

Our meta-analysis showed increased cortisol had a small predictive effect on onset or relapse and recurrence of MDD, which is in line with literature showing increased cortisol levels in MDD cross-sectionally [106, 107]. Yet, this effect disappeared when studies including baseline clinical diagnoses were excluded. Since increased cortisol is also a marker of stress [108], increased cortisol may be an indirect marker of previous stressful life events or stress induced by being ill. This underlines the importance of future research following *healthy samples* without subclinical depression longitudinally until a MDD diagnosis is established. Further, cortisol results were influenced by publication bias and study quality and the effect disappeared when outliers were removed or poor quality studies were removed. This underlines the need for high-quality prospective research on biomarkes for MDD.

Some limitations of the studies included and of the meta-analyses are noted. On a study level, poor quality studies were identified and small samples that develop MDD at follow-up were investigated. Neuroimaging studies use smaller samples than immunology and hormons studies. This limits the interpretation and generalization of findings for sample size topics. Further, we did not correct for multiple testing by applying $p = 0.05$ as threshold for significance. A correction would result in disappearance of the cortisol effect, indicating this may be a false positive. Based on our narrative synthesis heterogeneity of studies was visible and studies reporting no significant results were prominent, yet tend to not report sufficient data for inclusion in meta-analysis, resulting in a bias in the meta-analyses on significant effects. These limitations may have resulted in inflated odds ratios in our meta-analysis, and results should thus be interpreted with caution.

Overall, the findings of the current systematic review highlight the lack of prospective evidence for biomarkers as predictors of onset of MDD and relapse/recurrence. Our systematic search uncovers the causality gap that is present in biomarker research. It is striking not to find strong prospective evidence for any of the postulated biological theories. Thus, most of the leading hypotheses are based on results from cross-sectional research, treatment studies, symptomatology studies, or animal studies (e.g. [8, 12, 16, 18, 20]), which cannot determine causality [21]. Whether the observed changes in putative biomarker systems in MDD is a potential cause or consequence of depression thus remains unclear.

Our results, of course, do not indicate that there are *no* causal biomarkers, but highlight the dearth of *prospective evidence* that biomarkers explain onset, and relapse/recurrence of MDD. In addition, prospective evidence would suggest causality, yet it is only the *minimum* requirement for detecting causal relations. Manipulation studies should also be performed in order to demonstrate that alteration of one variable (biomarker) leads to the expected outcome (MDD). Indeed, experimental challenges including depletion studies, such as tryptophan depletion are available and have been shown to predict depressive relapse in certain circumstances [109]. Yet, a limitation of these studies is the temporary nature of the measured outcome (e.g., brief symptom reduction) and that common higher order biological (e.g. neuromodulatory) changes may also account for the differences in depletion responses [31, 109]. Combining different techniques from different biological levels may disentangle which factors are most directly causally linked to depression etiology. Future studies applying transcranial magnetic stimulation or other brain stimulation approaches to simulate symptoms/relapse may provide more insights into causal neuroimaging biomarkers [110]. It must be noted that we did not search for relatively newly identified biomarkers, such as fatty acids [111], which are not yet part of an established etiological theory. Thus, future research is necessary to investigate if novel biomarkers can predict MDD and replicate the current incidental findings.

Notwithstanding the overall lack of prospective evidence for leading biological models for onset, relapse and recurrence of MDD, future research may be directed to focus on potential predictive biomarkers identified in a small number of studies or showing inconsistent results. These were insula volume [36], thickness [51], and activity [99, 100] frontal brain activity [31, 39, 50, 76, 82], gastrointestinal sensitivity [40], norepinephrine [68], immunology markers [61, 87], androsterones [33], and oxidative stress markers [88]. Prospective research on these biomarkers investigating development of MDD from healthy samples is needed to

replicate these incidental finding and further investigate if predictive effects exist irrespectively of disease state. Indeed, there are indications that biomarkers may be causally involved, for example based on genetics research. Recent large consortium results (e.g. depression PGC [112]) have been successful in identifying genetic loci associated with depression. More importantly, depression is not a single gene disease but rather seems to be related to multiple genes in interaction with environmental factors, which lead to a spectrum of aversive outcomes, ranging from depressive symptoms to full-blown MDD [112]. The genetic loci identified explain only limited variance of depression (e.g. 2% genetic risk score [112] and mendelian randomization studies <1% [113]), whereas the heritability of MDD has been estimated at ~40% [114]. This suggests that MDD may be a more heterogeneous disorder both in etiology and pathophysiology. To unravel the biological mechanisms of MDD we therefore suggest to investiate interactions between biomarkers instead of investigating biomarkers separately for example in pathway or network approach.

In order to falsify biological theories for MDD better comparisons between or integration of studies is necessary. Open science initiatives could play a role in these efforts by enabling researchers to combine datasets over multiple cohorts (Consortia studies). However, it is noteworthy that there are large cohort samples available that allow prospective analysis on the clinical diagnosis MDD, yet clinical symptoms are more frequently investigated. In addition, baseline measurements where participants are healthy (before the development of MDD onset or relapse/recurrence) are frequently lacking in cohort studies. Further, investigating differential effects of onset versus relapse/recurrence is not common practice in biology research, whilst different mechanisms may underlie MDD onset versus maintenance. Future studies should separate samples with first onset from samples with previous episodes in order to investigate the differential mechanisms. Finally, given most theories on depression etiology include biological, psychological and social factors [115, 116], it is noteworthy that few studies have investigated combinations of these factors in a single study. Viewing depression from a more holistic perspective may help capture important interactions and improve prediction models.

Conclusion

This systematic search for prospective evidence for biomarkers of MDD revealed scarce prospective evidence for leading biological models. Prospective evidence for etiological involvement of gastrointestinal factors, neuroimaging, neurotrophic factors, neurotransmitters, hormones (other than cortisol), immunology and oxidative stress in MDD is lacking. Cortisol was found to be a predictor for onset/relapse/recurrence of MDD, but this effect was confounded by baseline clinical depression and quality of studies. Therefore, there is a need for high quality, prospective studies on the relative contribution of biomarkers (in combination with psychosocial factors) in order to disentangle the etiology of MDD and to better understand its clinical course.

Acknowledgements We would like to thank M. Brouwer, Z. Fu, A. Cramer, J. H. Ormel, P. Spinhoven, G. Siegle, S. D. Hollon, E. Holmes, C. Harmer, and M. van Hout for their input on the search process. In addition, we would like to thank the Clinical Psychology (Research) Master students from Utrecht University that worked on the project for their help with screening and data extraction on different biomarkers: I. Baltrusaityte, A. Mastora, A. Mroz, M. Barmpounis, J. Janssen, and B. Markovitch.

Funding Funding was partly received from the Netherlands Institute for Advanced Study in the Humanities and Social Sciences (NIAS), entitled "My Optimism Wears Heavy Boots: So much research, so few implications, towards 'patient-proof' empirical models and more effective interventions in mental health" (awarded to CB February–June 2017). This meta-analysis has been pre registered (PROSPERO 2017 CRD42017072990) and preliminary results were presented at the European congress of Psychology (2017 Amsterdam) and Society of Biological Psychiatry annual meeting (2018 New York). All authors have approved the content of the present form of the paper.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

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

References

1. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry. 2005;62:617–27.

2. Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B. Depression, chronic diseases, and decrements in health: results from the World Health Surveys. Lancet. 2007;370:851–8.
3. Hardeveld F, Spijker J, De Graaf R, Hendriks SM, Licht CMM, Nolen WA, et al. Recurrence of major depressive disorder across different treatment settings: results from the NESDA study. J Affect Disord. 2013;147:225–31.
4. Moffitt TE, Caspi A, Taylor A, Kokaua J, Milne BJ, Polanczyk G, et al. How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. Psychol Med. 2010;40:899–909.
5. Mueller TI, Leon AC, Keller MB, Solomon DA, Endicott J, Coryell W, et al. Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. Am J Psychiatry. 1999;156:1000–6.
6. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Lancet. 2018;391:1357–66.
7. Undurraga J, Baldessarini RJ. Randomized, placebo-controlled trials of antidepressants for acute major depression: thirty-year meta-analytic review. Neuropsychopharmacology. 2012;37:851–64.
8. Krishnan V, Nestler EJ. Linking molecules to mood: new insight into the biology of depression. Am J Psychiatry. 2010;167:1305–20.
9. Nestler EJ, Barrot M, DiLeone RJ, Eisch AJ, Gold SJ, Monteggia LM. Neurobiology of depression. Neuron. 2002;34:13–25.
10. Schlaepfer TE, Cohen MX, Frick C, Kosel M, Brodesser D, Axmacher N, et al. Deep brain stimulation to reward circuitry alleviates anhedonia in refractory major depression. Neuropsychopharmacology. 2008;33:368–77.
11. Price JL, Drevets WC. Neural circuits underlying the pathophysiology of mood disorders. Trends Cogn Sci (Regul Ed). 2012;16:61–71.
12. Wallace CJK, Milev R. The effects of probiotics on depressive symptoms in humans: a systematic review. Ann Gen Psychiatry. 2017;16:14.
13. Clapp M, Aurora N, Herrera L, Bhatia M, Wilen E, Wakefield S. Gut microbiota's effect on mental health: the gut-brain axis. Clin Pract. 2017;7:987.
14. Miller AH, Raison CL. The role of inflammation in depression: from evolutionary imperative to modern treatment target. Nat Rev Immunol. 2016;16:22–34.
15. Krystal JH, Sanacora G, Blumberg H, Anand A, Charney DS, Marek G, et al. Glutamate and GABA systems as targets for novel antidepressant and mood-stabilizing treatments. Mol Psychiatry. 2002;7:S71–80.
16. Andrews PW, Bharwani A, Lee KR, Fox M, Thomson JA. Is serotonin an upper or a downer? The evolution of the serotonergic system and its role in depression and the antidepressant response. Neurosci Biobehav Rev. 2015;51:164–88.
17. Brunoni AR, Lopes M, Fregni F. A systematic review and meta-analysis of clinical studies on major depression and BDNF levels: implications for the role of neuroplasticity in depression. Int J Neuropsychopharmacol. 2008;11:1169–80.
18. Duman RS, Li N. A neurotrophic hypothesis of depression: role of synaptogenesis in the actions of NMDA receptor antagonists. Philos Trans R Soc Lond B, Biol Sci. 2012;367:2475–84.
19. Holsboer F. The corticosteroid receptor hypothesis of depression. Neuropsychopharmacology. 2000;23:477–501.
20. Black CN, Bot M, Scheffer PG. Penninx BWJH. Oxidative stress in major depressive and anxiety disorders, and the association with antidepressant use; results from a large adult cohort. Psychol Med. 2017;47:936–48.
21. Mayeux R. Biomarkers: potential uses and limitations. NeuroRx. 2004;1:182–8.
22. Belmaker RH, Agam G. Major depressive disorder. N Engl J Med. 2008;358:55–68.
23. van Loo HM, Aggen SH, Gardner CO, Kendler KS. Multiple risk factors predict recurrence of major depressive disorder in women. J Affect Disord. 2015;180:52–61.
24. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6: e1000097.
25. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64:383–94.
26. Cuijpers P. Meta-analyses inmental health research: a practical guide. Pim Cuijpers Uitgeverij, Amsterdam, the Netherlands, 2016.
27. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21:1539–58.
28. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315:629–34.
29. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics. 2000;56:455–63.
30. Adam EK, Doane LD, Zinbarg RE, Mineka S, Craske MG, Griffith JW. Prospective prediction of major depressive disorder from cortisol awakening responses in adolescence. Psychoneuroendocrinology. 2010;35:921–31.*
31. Allen JJB, McKnight KM, Moreno FA, Demaree HA, Delgado PL. Alteration of frontal EEG asymmetry during tryptophan depletion predicts future depression. J Affect Disord. 2009;115: 189–95.*
32. Appelhof BC, Huyser J, Verweij M, Brouwer JP, van Dyck R, Fliers E, et al. Glucocorticoids and relapse of major depression (dexamethasone/corticotropin-releasing hormone test in relation to relapse of major depression). Biol Psychiatry. 2006;59:696–701.*
33. Asselmann E, Kische H, Haring R, Hertel J, Schmidt C-O, Nauck M, et al. Prospective associations of androgens and sex hormone-binding globulin with 12-month, lifetime and incident anxiety and depressive disorders in men and women from the general population. J Affect Disord. 2019;245:905–11.*
34. Aubry J-M, Gervasoni N, Osiek C, Perret G, Rossier MF, Bertschy G, et al. The DEX/CRH neuroendocrine test and the prediction of depressive relapse in remitted depressed outpatients. J Psychiatr Res. 2007;41:290–4.*
35. Banki CM, Karmacsi L, Bissette G, Nemeroff CB. CSF corticotropin-releasing hormone and somatostatin in major depression: response to antidepressant treatment and relapse. Eur Neuropsychopharmacol. 1992;2:107–13.*
36. Belden AC, Barch DM, Oakberg TJ, April LM, Harms MP, Botteron KN, et al. Anterior insula volume and guilt: neurobehavioral markers of recurrence after early childhood major depressive disorder. JAMA Psychiatry. 2015;72:40–8.*
37. Bockting CLH, Lok A, Visser I, Assies J, Koeter MW, Schene AH, et al. Lower cortisol levels predict recurrence in remitted patients with recurrent depression: a 5.5 year prospective study. Psychiatry Res. 2012;200:281–7.*
38. Bouhuys AL, Bos EH, Geerts E, van Os TWDP, Ormel J. The association between levels of cortisol secretion and fear perception in patients with remitted depression predicts recurrence. J Nerv Ment Dis. 2006;194:478–84.*
39. Bress JN, Foti D, Kotov R, Klein DN, Hajcak G. Blunted neural response to rewards prospectively predicts depression in adolescent girls. Psychophysiology. 2013;50:74–81.*

40. Campo JV, Dahl RE, Williamson DE, Birmaher B, Perel JM, Ryan ND. Gastrointestinal distress to serotonergic challenge: a risk marker for emotional disorder? J Am Acad Child Adolesc Psychiatry. 2003;42:1221–6.*
41. Carnegie R, Araya R, Ben-Shlomo Y, Glover V, O'Connor TG, O'Donnell KJ, et al. Cortisol awakening response and subsequent depression: prospective longitudinal study. Br J Psychiatry. 2014;204:137–43.*
42. Charles GA, Schittecatte M, Rush AJ, Panzer M, Wilmotte J. Persistent cortisol non-suppression after clinical recovery predicts symptomatic relapse in unipolar depression. J Affect Disord. 1989;17:271–8.*
43. Chocano-Bedoya PO, Mirzaei F, O'Reilly EJ, Lucas M, Okereke OI, Hu FB, et al. C-reactive protein, interleukin-6, soluble tumor necrosis factor α receptor 2 and incident clinical depression. J Affect Disord. 2014;163:25–32.*
44. Chopra KK, Segal ZV, Buis T, Kennedy SH, Levitan RD. Investigating associations between cortisol and cognitive reactivity to sad mood provocation and the prediction of relapse in remitted major depression. Asian J Psychiatr. 2008;1:33–6.*
45. Colich NL, Kircanski K, Foland-Ross LC, Gotlib IH. HPA-axis reactivity interacts with stage of pubertal development to predict the onset of depression. Psychoneuroendocrinology. 2015;55:94–101.*
46. Copeland WE, Shanahan L, Worthman C, Angold A, Costello EJ. Cumulative depression episodes predict later C-reactive protein levels: a prospective analysis. Biol Psychiatry. 2012;71:15–21.*
47. Coplan JD, Wolk SI, Goetz RR, Ryan ND, Dahl RE, Mann JJ, et al. Nocturnal growth hormone secretion studies in adolescents with or without major depression re-examined: integration of adult clinical follow-up data. Biol Psychiatry. 2000;47:594–604.*
48. Cosgriff JP, Abbott RM, Oakley-Browne MA, Joyce PR. Cortisol hypersecretion predicts early depressive relapse after recovery with electroconvulsive therapy. Biol Psychiatry. 1990; 28:1007–10.*
49. Davey CG, Whittle S, Harrison BJ, Simmons JG, Byrne ML, Schwartz OS, et al. Functional brain-imaging correlates of negative affectivity and the onset of first-episode depression. Psychol Med. 2015;45:1001–9.*
50. Farb NAS, Anderson AK, Bloch RT, Segal ZV. Mood-linked responses in medial prefrontal cortex predict relapse in patients with recurrent unipolar depression. Biol Psychiatry. 2011;70: 366–72.*
51. Foland-Ross LC, Sacchet MD, Prasad G, Gilbert B, Thompson PM, Gotlib IH. Cortical thickness predicts the first onset of major depression in adolescence. Int J Dev Neurosci. 2015;46:125–31.*
52. Franz B, Kupfer DJ, Miewald JM, Jarrett DB, Grochocinski VJ. Growth hormone secretion timing in depression: clinical outcome comparisons. Biol Psychiatry. 1995;38:720–9.*
53. Frodl T, Meisenzahl EM, Zetzsche T, Höhne T, Banac S, Schorr C, et al. Hippocampal and amygdala changes in patients with major depressive disorder and healthy controls during a 1-year follow-up. J Clin Psychiatry. 2004;65:492–9.*
54. Frodl T, Jäger M, Smajstrlova I, Born C, Bottlender R, Palladino T, et al. Effect of hippocampal and amygdala volumes on clinical outcomes in major depression: a 3-year prospective magnetic resonance imaging study. J Psychiatry Neurosci. 2008;33:423–30.*
55. Glaus J, von Känel R, Lasserre AM, lok MPF, Vandeleur CL, Castelao E, et al. Mood disorders and circulating levels of inflammatory markers in a longitudinal population-based study. Psychol Med. 2018;48:961–73.*
56. Goodyer IM, Herbert J, Tamplin A, Altham PM. First-episode major depression in adolescents. Affective, cognitive and endocrine characteristics of risk status and predictors of onset. Br J Psychiatry. 2000;176:142–9.*
57. Goodyer IM, Herbert J, Tamplin A, Altham PM. Recent life events, cortisol, dehydroepiandrosterone and the onset of major depression in high-risk adolescents. Br J Psychiatry. 2000;177:499–504.*
58. Goodyer IM, Bacon A, Ban M, Croudace T, Herbert J. Serotonin transporter genotype, morning cortisol and subsequent depression in adolescents. Br J Psychiatry. 2009;195:39–45.*
59. Goodyer IM, Croudace T, Dudbridge F, Ban M, Herbert J. Polymorphisms in BDNF (Val66Met) and 5-HTTLPR, morning cortisol and subsequent depression in at-risk adolescents. Br J Psychiatry. 2010;197:365–71.*
60. Grynderup MB, Kolstad HA, Mikkelsen S, Andersen JH, Bonde JP, Buttenschøn HN, et al. A two-year follow-up study of salivary cortisol concentration and the risk of depression. Psychoneuroendocrinology. 2013;38:2042–50.*
61. Haastrup E, Grau K, Eugen-Olsen J, Thorball C, Kessing LV, Ullum H. Soluble urokinase plasminogen activator receptor as a marker for use of antidepressants. PLoS ONE. 2014;9:e110555.*
62. Hardeveld F, Spijker J, Vreeburg SA, Graaf RD, Hendriks SM, Licht CMM, et al. Increased cortisol awakening response was associated with time to recurrence of major depressive disorder. Psychoneuroendocrinology. 2014;50:62–71.*
63. Harris TO, Borsanyi S, Messari S, Stanford K, Cleary SE, Shiers HM, et al. Morning cortisol as a risk factor for subsequent major depressive disorder in adult women. Br J Psychiatry. 2000;177:505–10.*
64. Hatzinger M, Hemmeter UM, Baumann K, Brand S, Holsboer-Trachsler E. The combined DEX-CRH test in treatment course and long-term outcome of major depression. J Psychiatr Res. 2002;36:287–97.*
65. Herbert J, Ban M, Brown GW, Harris TO, Ogilvie A, Uher R, et al. Interaction between the BDNF gene Val/66/Met polymorphism and morning cortisol levels as a predictor of depression in adult women. Br J Psychiatry. 2012;201:313–9.*
66. Jarrett DB, Kupfer DJ, Miewald JM, Grochocinski VJ, Franz B. Sleep-related growth hormone secretion is persistently suppressed in women with recurrent depression: a preliminary longitudinal analysis. J Psychiatr Res. 1994;28:211–23.*
67. Joffe RT, Marriott M. Thyroid hormone levels and recurrence of major depression. Am J Psychiatry. 2000;157:1689–91.*
68. Johnston TG, Kelly CB, Stevenson MR, Cooper SJ. Plasma norepinephrine and prediction of outcome in major depressive disorder. Biol Psychiatry. 1999;46:1253–8.*
69. Khandaker GM, Pearson RM, Zammit S, Lewis G, Jones PB. Association of serum interleukin 6 and C-reactive protein in childhood with depression and psychosis in young adult life: a population-based longitudinal study. JAMA Psychiatry. 2014; 71:1121–8.*
70. Kronmüller K-T, Pantel J, Köhler S, Victor D, Giesel F, Magnotta VA, et al. Hippocampal volume and 2-year outcome in depression. Br J Psychiatry. 2008;192:472–3.*
71. Langenecker SA, Jenkins LM, Stange JP, Chang Y-S, DelDonno SR, Bessette KL, et al. Cognitive control neuroimaging measures differentiate between those with and without future recurrence of depression. Neuroimage Clin. 2018;20:1001–9.*
72. LeMoult J, Ordaz SJ, Kircanski K, Singh MK, Gotlib IH. Predicting first onset of depression in young girls: Interaction of diurnal cortisol and negative life events. J Abnorm Psychol. 2015;124:850–9.*
73. Little K, Olsson CA, Youssef GJ, Whittle S, Simmons JG, Yücel M, et al. Linking the serotonin transporter gene, family environments, hippocampal volume and depression onset: A prospective imaging gene × environment analysis. J Abnorm Psychol. 2015;124:834–49.*
74. Little K, Olsson CA, Whittle S, Youssef GJ, Byrne ML, Simmons JG, et al. Association between serotonin transporter

genotype, brain structure and adolescent-onset major depressive disorder: a longitudinal prospective study. Transl Psychiatry. 2014;4:e445.*
75. Lok A, Mocking RJT, Ruhé HG, Visser I, Koeter MWJ, Assies J, et al. Longitudinal hypothalamic-pituitary-adrenal axis trait and state effects in recurrent depression. Psychoneuroendocrinology. 2012;37:892–902.*
76. Lythe KE, Moll J, Gethin JA, Workman CI, Green S, Lambon Ralph MA, et al. Self-blame-selective hyperconnectivity between anterior temporal and subgenual cortices and prediction of recurrent depressive episodes. JAMA Psychiatry. 2015;72:1119–26.*
77. Macoveanu J, Baaré W, Madsen KH, Kessing LV, Siebner HR, Vinberg M. Risk for affective disorders is associated with greater prefrontal gray matter volumes: a prospective longitudinal study. Neuroimage Clin. 2018;17:786–93.*
78. Mander AJ, Rubin RT, Copolov DL, Poland RE. The predictive power of the salivary cortisol dexamethasone suppression test for three-year outcome in major depressive illness. J Psychiatr Res. 1989;23:151–6.*
79. Mocking RJT, Pellikaan CM, Lok A, Assies J, Ruhé HG, Koeter MW, et al. DHEAS and cortisol/DHEAS-ratio in recurrent depression: state, or trait predicting 10-year recurrence? Psychoneuroendocrinology. 2015;59:91–101.*
80. Morris MC, Rao U, Garber J. Cortisol responses to psychosocial stress predict depression trajectories: social-evaluative threat and prior depressive episodes as moderators. J Affect Disord. 2012;143:223–30.*
81. Nickson T, Chan SWY, Papmeyer M, Romaniuk L, Macdonald A, Stewart T, et al. Prospective longitudinal voxel-based morphometry study of major depressive disorder in young individuals at high familial risk. Psychol Med. 2016;46:2351–61.*
82. Nixon NL, Liddle PF, Nixon E, Worwood G, Liotti M, Palaniyappan L. Biological vulnerability to depression: linked structural and functional brain network findings. Br J Psychiatry. 2014;204:283–9.*
83. Nusslock R, Shackman AJ, Harmon-Jones E, Alloy LB, Coan JA, Abramson LY. Cognitive vulnerability and frontal brain asymmetry: common predictors of first prospective depressive episode. J Abnorm Psychol. 2011;120:497–503.*
84. Owashi T, Otsubo T, Oshima A, Nakagome K, Higuchi T, Kamijima K. Relationships of DEX/CRH and GHRH test results to the outcome of depression–preliminary results suggest the GHRH test may predict relapse after discharge. J Psychiatr Res. 2008;42:356–64.*
85. Papmeyer M, Giles S, Sussmann JE, Kielty S, Stewart T, Lawrie SM, et al. Cortical thickness in individuals at high familial risk of mood disorders as they develop major depressive disorder. Biol Psychiatry. 2015;78:58–66.*
86. Papmeyer M, Sussmann JE, Stewart T, Giles S, Centola JG, Zannias V, et al. Prospective longitudinal study of subcortical brain volumes in individuals at high familial risk of mood disorders with or without subsequent onset of depression. Psychiatry Res Neuroimaging. 2016;248:119–25.*
87. Pasco JA, Nicholson GC, Williams LJ, Jacka FN, Henry MJ, Kotowicz MA, et al. Association of high-sensitivity C-reactive protein with de novo major depression. Br J Psychiatry. 2010;197:372–7.*
88. Pasquali MA, Harlow BL, Soares CN, Otto MW, Cohen LS, Minuzzi L, et al. A longitudinal study of neurotrophic, oxidative, and inflammatory markers in first-onset depression in midlife women. Eur Arch Psychiatry Clin Neurosci. 2017;268:771–81.*
89. Pintor L, Torres X, Navarro V, Martinez de Osaba MAJ, Matrai S, Gastó C. Prediction of relapse in melancholic depressive patients in a 2-year follow-up study with corticotropin releasing factor test. Prog Neuropsychopharmacol Biol Psychiatry. 2009;33:463–9.*
90. Pintor L, Torres X, Bailles E, Navarro V, de Osaba MJM, Belmonte A, et al. CRF test in melancholic depressive patients with partial versus complete relapses: a 2-year follow-up study. Nord J Psychiatry. 2013;67:177–84.*
91. Rao U, Chen L-A, Bidesi AS, Shad MU, Thomas MA, Hammen CL. Hippocampal changes associated with early-life adversity and vulnerability to depression. Biol Psychiatry. 2010;67:357–64.*
92. Rao U, Hammen CL, Poland RE. Risk markers for depression in adolescents: sleep and HPA measures. Neuropsychopharmacology. 2009;34:1936–45.*
93. Rao U, Hammen CL, Poland RE. Longitudinal course of adolescent depression: neuroendocrine and psychosocial predictors. J Am Acad Child Adolesc Psychiatry. 2010;49:141–51.*
94. Rudaz DA, Vandeleur CL, Gebreab SZ, Gholam-Rezaee M, Strippoli M-PF, Lasserre AM, et al. Partially distinct combinations of psychological, metabolic and inflammatory risk factors are prospectively associated with the onset of the subtypes of Major Depressive Disorder in midlife. J Affect Disord. 2017;222:195–203.*
95. Serra-Blasco M, de Diego-Adeliño J, Vives-Gilabert Y, Trujols J, Puigdemont D, Carceller-Sindreu M, et al. Naturalistic course of major depressive disorder predicted by clinical and structural neuroimaging data: a 5-year follow-up. Depress Anxiety. 2016;33:1055–64.*
96. Tsuru J, Ishitobi Y, Ninomiya T, Kanehisa M, Imanaga J, Inoue A, et al. The thyrotropin-releasing hormone test may predict recurrence of clinical depression within ten years after discharge. Neuro Endocrinol Lett. 2013;34:409–17.*
97. Vinberg M, Miskowiak K, Kessing LV. Brain Derived Neurotrophic Factor (BDNF) levels as a possible predictor of psychopathology in healthy twins at high and low risk for affective disorder. Psychoneuroendocrinology. 2014;39:179–83.*
98. Vrshek-Schallhorn S, Doane LD, Mineka S, Zinbarg RE, Craske MG, Adam EK. The cortisol awakening response predicts major depression: predictive stability over a 4-year follow-up and effect of depression history. Psychol Med. 2013;43:483–93.*
99. Whalley HC, Sussmann JE, Romaniuk L, Stewart T, Papmeyer M, Sprooten E, et al. Prediction of depression in individuals at high familial risk of mood disorders using functional magnetic resonance imaging. PLoS ONE. 2013;8:e57357.*
100. Whalley HC, Sussmann JE, Romaniuk L, Stewart T, Kielty S, Lawrie SM, et al. Dysfunction of emotional brain systems in individuals at high risk of mood disorder with depression and predictive features prior to illness. Psychol Med. 2015;45:1207–18.*
101. Workman CI, Lythe KE, McKie S, Moll J, Gethin JA, Deakin JFW, et al. A novel resting-state functional magnetic resonance imaging signature of resilience to recurrent depression. Psychol Med. 2017;47:597–607.*
102. Zimmerman M, Coryell W, Pfohl B. Prognostic validity of the dexamethasone suppression test: results of a six-month prospective follow-up. The American Journal of Psychiatry 1987;144:212–214.*
103. Zobel AW, Yassouridis A, Frieboes RM, Holsboer F. Prediction of medium-term outcome by cortisol response to the combined dexamethasone-CRH test in patients with remitted depression. Am J Psychiatry. 1999;156:949–51.*
104. Zobel AW, Nickel T, Sonntag A, Uhr M, Holsboer F, Ising M. Cortisol response in the combined dexamethasone/CRH test as predictor of relapse in patients with remitted depression. a prospective study. J Psychiatr Res. 2001;35:83–94.*
105. Müller VI, Cieslik EC, Laird AR, Fox PT, Radua J, Mataix-Cols D, et al. Ten simple rules for neuroimaging meta-analysis. Neurosci Biobehav Rev. 2018;84:151–61.
106. Geerlings MI, Gerritsen L. Reply to: late-life depression, cortisol, and the hippocampus: on the need to consider depressive,

hippocampal, and pharmacological complexities. Biol Psychiatry. 2018;83:e25.
107. Stetler C, Miller GE. Depression and hypothalamic-pituitary-adrenal activation: a quantitative summary of four decades of research. Psychosom Med. 2011;73:114–26.
108. Miller GE, Chen E, Zhou ES. If it goes up, must it come down? Chronic stress and the hypothalamic-pituitary-adrenocortical axis in humans. Psychol Bull. 2007;133:25–45.
109. Ruhé HG, Mason NS, Schene AH. Mood is indirectly related to serotonin, norepinephrine and dopamine levels in humans: a meta-analysis of monoamine depletion studies. Mol Psychiatry. 2007;12:331–59.
110. Etkin A. Addressing the causality gap in human psychiatric neuroscience. JAMA Psychiatry. 2018;75:3–4.
111. Lin P-Y, Huang S-Y, Su K-P. A meta-analytic review of polyunsaturated fatty acid compositions in patients with depression. Biol Psychiatry. 2010;68:140–7.
112. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nat Genet. 2018;50:668–81.
113. Köhler CA, Evangelou E, Stubbs B, Solmi M, Veronese N, Belbasis L, et al. Mapping risk factors for depression across the lifespan: an umbrella review of evidence from meta-analyses and Mendelian randomization studies. J Psychiatr Res. 2018;103:189–207.
114. Sullivan PF, Neale MC, Kendler KS. Genetic epidemiology of major depression: review and meta-analysis. Am J Psychiatry. 2000;157:1552–62.
115. Beck AT, Bredemeier K. A unified model of depression: integrating clinical, cognitive, biological, and evolutionary perspectives. Clin Psychological Sci. 2016;4:596–619.
116. Marlies E. Brouwer, Alishia D. Williams, Mitzy Kennis, Zhongfang Fu, Nicola S. Klein, Pim Cuijpers, Claudi L.H. Bockting. Psychological theories of depressive relapse and recurrence: A systematic review and meta-analysis of prospective studies. Clinical Psychology Review:101773. (2019).
117. Nixon NL, Liddle PF, Worwood G, Liotti M, Nixon E. Prefrontal cortex function in remitted major depressive disorder. Psychological Medicine. 2013;43:1219–1230.
118. Mocking, R.J.T., Ruhé, H.G., Assies, J., et al. Relationship between the hypothalamic-pituitary-adrenal-axis and fatty acid metabolism in recurrent depression. Psychoneuroendocrinology. 2013;38:1607–1617.
*(References denoted by an asterisk are included in the narrative synthesis and/or meta-analyses).