Journal of Affective Disorders 393 (2026) 120387

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Tracking positive emotions in the course of adult depression: A systematic scoping review of longitudinal studies

S. Dupuis-Lesavre [a], C. Mebazaa [b,c], C. Veal [b,c], A. Ribeiro [b,c], I. Boutron [b,c], K. Krause [b,c], Y. Morvan [a,d], A. Chevance [b,c,*]

[a] Université Paris Nanterre, UFR SPSE, laboratoire CLIPSYD, EA4430, Nanterre, France
[b] Université de Paris Cité and Université Sorbonne Paris Nord, Inserm, INRAE, Center for Research in Epidemiology and StatisticS (CRESS), INSERM UMR 1153, F-75004, Paris, France
[c] Centre d'Épidémiologie Clinique, AP-HP, Hôpital Hôtel Dieu, F-75004, Paris, France
[d] Center for research in epidemiology and population health (CESP), National Institute of Health and Medical Research (INSERM) U1018, Paris-Saclay University, Villejuif, France

ARTICLE INFO

Keywords:
Depression
Bipolar disorder
Positive emotions
Ecological momentary assessment
Systematic review

ABSTRACT

Background: Research on depression has mostly focused on negative emotion with limited attention to the longitudinal evolution of positive emotions. There is a lack of clarity on the methods commonly used for tracking positive emotions in depression research and clinical practice. This methodological scoping review aimed to describe measurement procedures and instruments and appraise the quality of patient-reported outcome measures (PROMs).

Methods: PubMed, PSYCinfo, and EMBASE were systematically searched for longitudinal observational studies that measured positive emotions in adults with depression. We extracted data on measurement procedures and instruments, and conducted a COSMIN appraisal of instrument development and content validity for identified PROMs.

Results: As of April 16, 2024, 19 studies with a median sample size of 60 patients were included. Ecological Momentary Assessment (EMA) was employed by 81% of studies, though methods varied in frequency (1–10 measures/day), duration (3–50 days), prompts, data collection tools (e.g., booklets, apps), and incentives to minimize missing data. Positive emotions were measured using lists of adjectives (1–13 per list), with a total of 37 unique adjectives, some not strictly emotional but cognitive or behavioral. Two PROMs, the State-Trait Depression Adjective Checklist (ST-DACL) and the Positive and Negative Affect Schedule (PANAS), were used, though intended to assess moods rather than emotions. COSMIN ratings deemed the PROMs as doubtful due to poor reporting.

Conclusions: Establishing reporting guidelines for EMA studies and consensus on standardized procedures for monitoring positive emotions could enhance research synthesis and aid clinical assessments of depression's course.

1. Introduction

Emotions are defined in biomedical and psychological research as a "complex reaction pattern involving experiential, behavioral, and psychological elements" (APA. American Psychological Association, 2018). Besides arousal, the valence of emotions, *ie* the pleasantness or unpleasantness experienced by the individuals, is considered as a core dimension of emotions and is commonly used to classify them as *positive* or *negative* (Barrett et al., 2007; Snyder et al., 2021). In this review, we distinguish between related but distinct constructs. We use *positive emotions* to refer to valenced, discrete, and short-lived states (e.g., joy, enthusiasm, contentment). In contrast, *positive affect* denotes a higher-order dimensional tone of experience, while *mood* refers to longer-lasting, more diffuse affective contexts. In the remainder of the manuscript, we use the term "positive emotions" when reporting the terminology of the original authors, while applying our own conceptual framework in the analysis.

Emotional symptoms are core elements to the lived experience of

* Corresponding author at: Hôpital Hôtel Dieu, Centre d'Épidémiologie Clinique, 1 Place du Parvis Notre-Dame, Paris, 75004, France
E-mail address: astrid.chevance@aphp.fr (A. Chevance).

https://doi.org/10.1016/j.jad.2025.120387
Received 29 March 2025; Received in revised form 6 September 2025; Accepted 6 October 2025
Available online 10 October 2025

depression, with an excess of negative emotions and the diminution of positive emotions being cited as difficult to live with (Chevance et al., 2020). Moreover, emotional blunting, of both positive and negative emotions, as a troublesome adverse event of antidepressants, is reported by nearly half of those undergoing monoaminergic antidepressant treatment (Chevance et al., 2022; Goodwin et al., 2017).

The psychological model of depression acknowledges the role of positive and negative emotions. For example, the tripartite model of anxiety and depression by Clark and Watson (Clark and Watson, 1991) considers the absence of positive emotions as a key distinctive feature of depression for taxonomic purposes, whereas negative emotions are nonspecific.

A number of studies and evidence syntheses have investigated the prevalence and interplay of positive emotions in depressed people in cross-sectional observational studies, but fewer studies have focused on understanding the longitudinal course of positive emotions in depression under real-life conditions (de Vos et al., 2017; Panaite et al., 2021; Panaite et al., 2020). A better knowledge of the evolution of positive emotions in depression may be of particular interest for a range of clinical applications, such as refining diagnostic tools, prognosis, predicting treatment outcomes, and monitoring treatment efficacy. In fact, some authors concluded that lower daily positive affects predict higher depressive symptoms at 6 months, independently from the level of negative affect (Panaite et al., 2020). However, the conclusion of this study might be limited by the reduced sample size from the same community setting. Other authors showed that variability, instability, and inertia of positive affects were related to cognitive behavioral therapy outcomes, whereas for negative affects, only instability was (Bosley et al., 2019).

To date, monitoring the course of positive emotions in longitudinal studies remains uncertain due to the methodological challenge of measuring the normal fluctuation of emotions within a day, which can vary notably depending on the events encountered by the person (Barrett et al., 2007). Hence, measuring positive emotions at one single time point or outside the context of daily life, such as during interviews at a hospital or in a research center, may only reflect a general or broad overview with a lack of temporal specificity. Technological advances such as Ecological Momentary Assessment (EMA) - sometimes also called Experience Sampling Method (ESM)- facilitate more frequent and more ecologically embedded data collection (Hamaker and Wichers, 2017; Shiffman et al., 2008). Namely, EMA involves the repeated sampling of current experiences, behaviors, thoughts, and emotions in real-time, in turn minimizing recall bias and maximizing real-world validity (Shiffman et al., 2008). However, what kind of positive emotions are measured, at what frequency, and how they are measured in EMA longitudinal studies, remains unclear. Furthermore, the use of heterogeneous measures across studies could compromise the combination of their results in meta-analysis. For instance, a repository of EEMA items used in various populations (with or without mental disorders) and within various study designs (interventional, observational) showed the use of 29 unique adjectives to measure “positive affects”, of which half were measured by three studies or less (Kirtley et al., 2018) https://esmitemrepositoryinfo.com/.

This systematic scoping review has two aims. Endorsing a meta-research approach, the first aim is to describe the procedures and measurement instruments used by researchers to assess positive emotions in longitudinal observational studies assessing the course of depression in adults. The second aim is to assess the methodological quality of the development and the content validity of any patient-reported outcome measures (PROMs) used to measure positive emotions. This assessment will follow the process and criteria outlined by the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) guidelines.

2. Method

We conducted a systematic scoping review of the methods used to measure the course of positive emotions in longitudinal non-interventional studies of adult depression. In addition, we evaluated the development and content validity of the patient-reported outcome measures (PROM) identified by our review, using the COSMIN methodology (Terwee et al., 2018; Mokkink et al., 2018; Tricco et al., 2018). This review is compliant with the Joanna Briggs Institute methodology and the PRISMA guidelines for scoping reviews (Tricco et al., 2018). The protocol was registered on Zenodo (10.5281/zenodo.8146118), with the final version uploaded on July 14, 2023, following a pilot study conducted to test the search strategy and develop the data charting form.

2.1. Eligibility criteria of the studies included in the review

We included peer-reviewed scientific publications of longitudinal observational studies, with no minimum number of participants, whose primary or secondary objective was to investigate the course of positive emotions in adults (>18 years old) with depression. We included studies who reported populations with unipolar and/or bipolar depression, defined as either: one of the diagnoses of the Depressive Disorders or Bipolar and Related Disorders section of the Diagnostic and Statistical Manual of Mental Disorders (DSM) – 4th or 5th edition; one of the diagnoses of the Depressive Disorders or Bipolar or related disorders section of the International Classification of Diseases (ICD-10); scoring above a predetermined threshold, as defined by the study, on a depression measurement instrument (e.g., the Beck Depression Inventory, the Montgomery-Åsberg Depression Rating Scale); or a patient-reported diagnosis of depression.

To be included, studies had to assess at least one « positive emotion » on at least two occasions over a period longer than 24 hours, using a standardized tool that the authors described as measuring « positive emotions ». For publications derived from the same cohort (e.g., NESDA-EMAA, MOOVD), we treated each publication as a distinct study when it addressed a different research aim or applied different inclusion/exclusion criteria. We excluded studies on anhedonia, i.e., the global lack of pleasure, except if they specifically monitored positive emotions. We excluded studies on quality of life, life satisfaction, and well-being, since these constructs refer to an experience exceeding the feeling of positive emotions (Snyder et al., 2021). Moreover, there already are specific systematic reviews on these constructs (Dronavalli and Thompson, 2015; Linton et al., 2016; Moskowitz et al., 2021). Finally, we excluded letters, conference abstracts, posters, studies only involving remitted patients, studies assessing positive emotions exclusively via qualitative methods (i.e., semi-structured interviews, focus groups), and interventional studies defined as any studies that aimed to measure the efficacy of a treatment (e.g., randomized controlled trials, pre/post studies).

2.2. Information sources and search strategy

Given that dedicated keywords or search terms do not reference positive emotions in any of the databases of interest, we developed a search strategy involving four clinicians/researchers (AC, SDL, CM, and AR) and one professional librarian of Université Paris-Cité (Mr Colin Sidre) to identify a list of relevant keywords. The full search strategies are provided in Supplementary Material 1. The final search equation was constructed iteratively through a series of steps. First, we searched systematic reviews covering the topic of positive emotions on PubMed on December 19, 2022, with the keywords *positive emotions* OR *positive affects* AND *systematic review*. Three relevant systematic reviews were identified (Bassett et al., 2019; DuBois et al., 2015; Ong et al., 2017). Additionally, we identified two prominent books in the field of positive emotions (Snyder et al., 2021; Fredrickson and Cohn, 2008). From these sources, we extracted the names of the positive emotions cited and the

different measurement instruments used to evaluate them. Then, we formed a consensus on which keywords should be selected for our search strategy. We excluded ambiguous words that can lead to both positive and negative feelings, such as "surprise". In addition to the generic terms used to designate positive emotions (i.e. positive emotions, positive affect, positive mood, positive valence, hedonic), we included a list of specific positive emotions (i.e. happiness, pleasure, enjoyment, joyful, content, amusement, enthusiast, cheerful, humor, excitement, optimism, hope, love, awe, serenity, serene, gratitude, grateful, calm, peaceful, harmony, calmness, elation, elated, happy, euphoric). Finally, we combined this list with a search block of terms dedicated to depression and a search block of terms dedicated to longitudinal studies adapted from the Biomedical Information Group of the Netherlands (https://blocks.bmi-online.nl/).

As recommended by the Cochrane methodology for systematic reviews, we searched three databases relevant to psychology and psychiatry: PubMed, PsycINFO, and Embase from inception to the date of extraction (April 16, 2024).

All references identified by the search strategy of the three databases were then uploaded to the COVIDENCE software data manager and screening tool.

2.3. Selection of the relevant scientific publications

Using the eligibility criteria, each title and abstract was independently screened by two of the three reviewers (SDL, CM, and AC). All conflicts were discussed until a consensus was reached. When some words could simultaneously refer to positive emotions or mood or personality traits, depending on the context of use (e.g., "optimism" or "elation"), the reviewers discussed each specific case and decided on what it refers to through semantic analysis of its context of use. Then, the full text of each of the approved studies was independently screened by the three reviewers using the same eligibility criteria.

2.4. Data extraction

For all included studies, one researcher (either SDL or CM) extracted the data, which was then double-checked by the other. All extracted data were subsequently reviewed a third time by AC. Discrepancies were resolved by discussing with other members of the research team. The data extracted to describe the sample of studies included: the publications' metadata (e.g., title, date of publication), the objective of the study, and the population included (e.g., eligibility criteria, country, sample size, gender proportion). In instances where the study took place in a cohort, we extracted the cohort's characteristics (e.g., name, date of creation, setting, inclusion and exclusion criteria, type of depression, and sample size). The data extracted to describe the procedures and measurement instruments used by researchers to assess "positive emotions" included: the design used to assess the course of "positive emotions" (e.g., EMA), the collection mode (e.g., type of electronic device), the measurement instrument used to assess "positive emotions", the name of the "positive emotions" assessed (and negative emotions if any), the number of times measures were repeated, the duration of the follow-up, and the use of incentives to limit missing data). We considered all measurement instruments used to measure "positive emotions", as stated by the authors, whether they were developed to specifically measure "positive emotions" or not.

2.5. Analysis

We used descriptive statistics to describe the included studies. To address the first aim, we reported the frequency of each procedure used to assess "positive emotions". For instance, we calculated the frequencies and proportions of each "positive emotion" measured across the sample of studies included in the review.

To address the second aim, we evaluated the development and content validity of the patient-reported outcome measures (PROM) identified, using the COSMIN guidelines (Terwee et al., 2017). We replicated the same method we previously conducted for another construct (mental pain) (Charvet et al., 2022). We included the different versions of a single tool, if any (e.g., shorter scales), as long as they had at least one development and validation publication.

We identified the relevant publication to evaluate the development and content validity of each PROM with two processes. First, we extracted the development and validation studies cited by the studies retrieved by our review. Second, we emailed the corresponding author of each instrument to request all development/validation materials (published and unpublished studies). Then, two investigators (SD and CM) independently rated the quality of the PROM development and the content validity. A third researcher (AC) checked the ratings. Disagreements, if any, were resolved by consensus with the other researchers of the team.

In brief, 35 features were assessed for the development, across two categories: (1) "design" (e.g., conceptual framework, target population description, qualitative data methods) and (2) "cognitive interview study or pilot test" (e.g., patient feedback on measure comprehensiveness). Features were rated as 'very good', 'adequate', 'doubtful', 'inadequate', or 'not applicable'. Following the COSMIN recommendation that 'poor methodological aspects of a study cannot be compensated by [different] good aspects', the overall rating of each primary category is the lowest score of the features. However, COSMIN differentiates 'fatal flaws', which lead automatically to the rating of the entire category as 'inadequate', from other flaws that are rated adequate or doubtful, to lower their impact in global scoring. For instance, not using an appropriate qualitative data collection method to identify relevant items for a new PROM leads to the rating of the overarching 'design' category as 'inadequate'.

The quality of content validity studies was assessed on relevance, comprehensiveness, and comprehensibility, across 31 features classified in 5 main categories: (1) 'asking patients about the relevance of PROM items', (2) 'asking patients about comprehensiveness', (3) 'asking patients about comprehensibility', (4) 'asking professionals about relevance' and (5) 'asking professionals about comprehensiveness'. Details of the 31 features can be found in box 2 of the 'COSMIN methodology for content validity' user manual (Terwee et al., 2018). These features are clarified by questions such as 'Was an appropriate method used for assessing the relevance/comprehensiveness/comprehensibility of the PROM?' 'Were skilled moderators used?' 'Was each item tested in an adequate number of patients?' 'Were professionals from all relevant disciplines included?' Each of the 31 domains can be rated 'very good', 'adequate', 'doubtful', 'inadequate' or 'not applicable'. Like the assessment of development, the overall rating is based on the lowest score of the features. The investigators reported no conflict of interest with any of the measures.

3. Findings

The search conducted on April 16, 2024 retrieved 5640 unique references (Fig. 1). After screening on title and abstract, the full texts of 93 references were assessed for eligibility. In total, 19 research articles published between 1977 and 2023 whose objective was to evaluate the course of "positive emotions" in depressive disorders and/or in bipolar depression were included (a list is available in Supplementary Material 2). Of the 19 publications included, 16 unique studies were identified and are described in Table 1 and Supplementary Material 3 (Table A and B).

Among the 16 studies, three relied on cohort data: the Netherlands Study of Depression and Anxiety Actigraphy sub-study (NESDA-EMAA), the Midlife Development in the United States (MIDUS) study, and the East Flanders Prospective Twin Survey (EFPTS). Overall, the median number of participants included was 108 (min= 39, max= 5898). Regarding participants with ongoing or past depression, the median

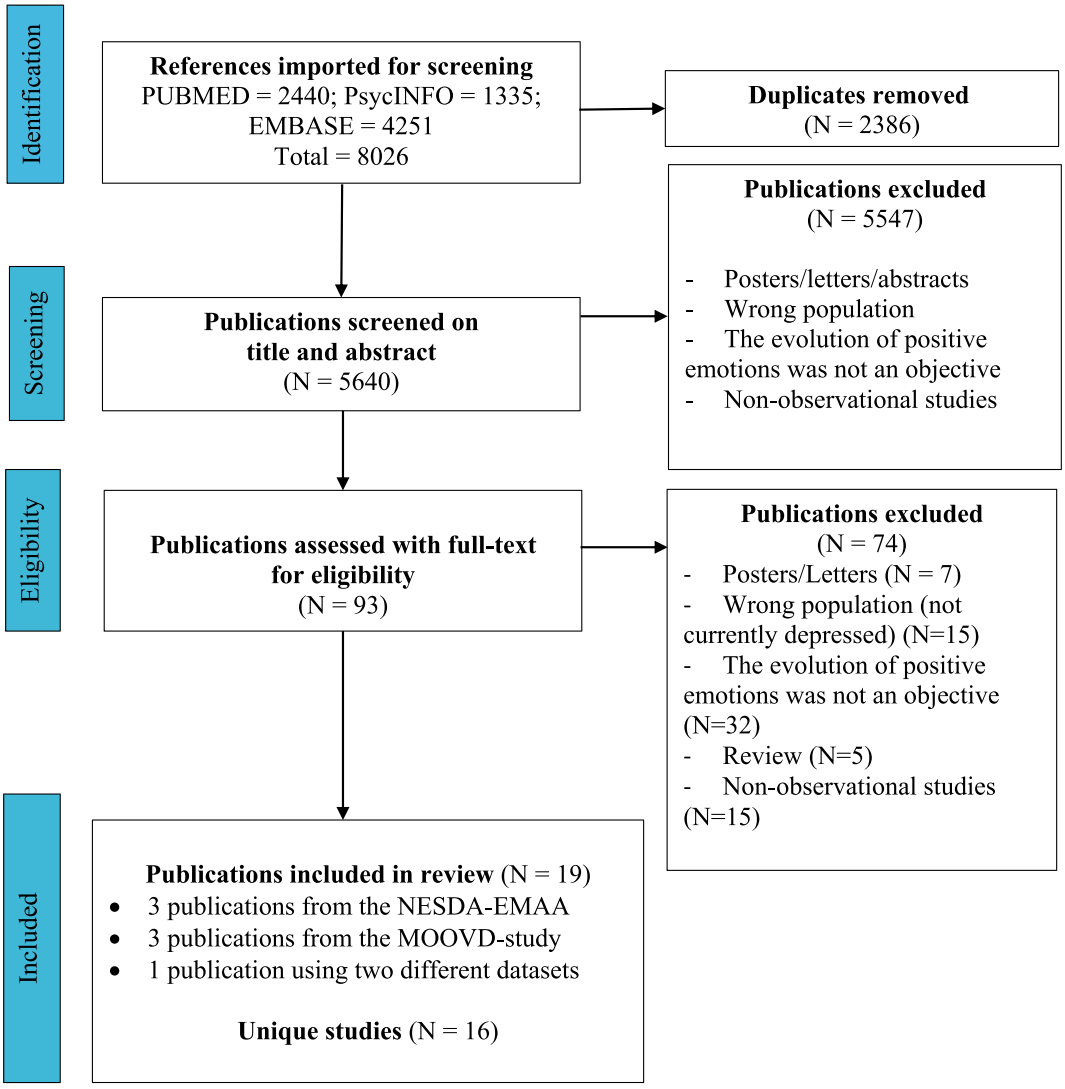

Fig. 1. Flowchart of the systematic scoping review.

number of participants with depression was 60 (min= 11, max= 2949). Three studies included participants with both bipolar or unipolar depression (Baik and Newman, 2023; Heininga et al., 2019; Hornstra and Klassen, 1977). More than half of the studies were conducted in the United States of America (9/16), and a quarter were conducted in the Netherlands (4/16). Three studies involved exclusively females (Kircanski et al., 2018; Maddever and Calhoun, 1986; Minaeva et al., 2021).

3.1. Procedures used to measure the course of "positive emotions" across the studies

Regarding the research design used to collect data on "positive emotions", 81% of the studies (13/16) used the Ecological Momentary Assessment (EMA) (Supplementary Material 3, Table A). The three remaining studies, which were published before 2013, used a collection of two or three measures at one-year and two-year intervals (Hornstra and Klassen, 1977; Brown, 2007; García-Peña et al., 2013). Nearly all EMA studies (12/13) used data collection modes which prompted participants with electronic devices (e.g., smartphone, electronic diary, wristwatch) using fixed or semi-random prompts (Supplementary Material 3, Table A). Patients then had to report their current "positive emotions" through self-reported paper and pencil research booklets (3/13) or electronic devices such as an electronic diary (4/13) or smartphone app (5/13) (Table 1). The remaining EMA study, which was published in 1986, used a daily phone call at a fixed hour in the evening (Maddever and Calhoun, 1986). Regarding the number of times measures were repeated in EMA studies, with measures of "positive emotions" being collected from 1 to 10 times a day (median 8) for a period of 3 to 50 consecutive days (median 8) (Supplementary Material 3, Table A). Of note, five studies reported having used incentives to limit missing data: 2 used monetary or voucher compensation, 2 sent personalized feedback on the emotional pattern of the participants after the study completion, and 1 involving students provided course credits.

3.2. Measurement instruments used to assess "positive emotions" across the studies

All of the studies used a list of adjectives which the authors referred to as "positive emotions" (Supplementary material 3, Table B). All but two also used a list of "negative emotions" (Supplementary material 3, Table B). These lists of adjectives were either standardized tools (4 studies), or derived from standardized tools (7 studies), or *ad hoc* lists specifically developed for the study (5 studies) (Table 1).

Among the four studies that used standardized tools, two used the State-Trait Depression Adjectives Checklist (ST-DACL) and two used the Positive and Negative Affect Schedule (PANAS) (Table 1). Their development is extensively described in Supplementary material 4.

Table 1
Sample, procedures, and measurement instruments used to evaluate positive emotions (N=16 unique studies)

Publications	Sample size[a]	Sampling method	Measurement Instrument (Number of Positive Emotions); Response anchor	Study duration
Hornstra, 1977	N= ?/342 (USA)	Interview; 1/year	ST-DACL (12 PE); Checklist	12 months
Maddever, 1986	N= 26/39 (USA)	EMA; 1/day	ST-DACL (12 PE); Checklist	14 days
Peeters, 2006	N= 47/86 (Netherlands)	EMA; 10/day	*Ad hoc* list of adjectives (8 PE); Intensity	6 days
Brown, 2007	N= 218/606 (USA)	Interview; 1/year	PANAS (10 PE); Intensity	24 months
Garcia-Pena 2013	N= 2949/5898 (USA)	Interview; 1/year	4 items from the CESD (4 PE); Frequency	24 months
Kircanski, 2017	N= 36/70 (USA)	EMA; 8/day	*Ad hoc* list derived from PANAS-X (4 PE) and Ekman's basic emotions (1PE) (Ekman et al., 1972); Intensity	8 days
Mood and Movement in Daily Life - MOOVD Study (3 publications)	N= 27/54 (Netherlands) N= 27/54 (Netherlands) N=10/20 (Netherlands)	EMA; 3/day	*Ad hoc* list derived from PANAS-X (5 PE), and some additional adjectives (2 PE); Intensity	30 days
Heininga, 2019	N= 47/87 (Belgium)	EMA; 10/day	*Ad hoc* list derived from PANAS-X (3 PE); Intensity	7 days
Panaite, 2020	N= 60/98 (USA)	EMA; 10/day	*Ad hoc* list derived from PANAS-X (5 PE), and some additional adjectives (2 PE); Intensity	3 days
Substudy EMA and Actigraphy of the NESDA-EMAA (3 publications)	N= 273/365 (Netherlands) N= 269/359 (Netherlands) N= 221/279 (Netherlands)	EMA; 5/day	*Ad hoc* list of adjectives (6 PE); Intensity	2 weeks
Shin, 2021 (2 studies)	N= 61/119 (USA)	EMA; 9/day	*Ad hoc* list of adjectives inspired from the circumplex model of affect by Russel (4 PE) (Russell, 1980; Russell and Barrett, 1999); Intensity	8 days
	N= 97/169 (USA)	EMA; 1/day	PANAS (10 PE); Intensity	50 days
Minaeva, 2021	N= 108/579 (Belgium)	EMA; 10/day	*Ad hoc* list of adjectives (4 PE); Intensity	5 days
Panaite, 2021	N= 121/839 (USA)	EMA; 1/day	*Ad hoc* list derived from PANAS-X (8 PE) and some additional adjectives (5 PE); Frequency	8 days
Lucht, 2022	N= 74/74 (Germany)	EMA; 10/day	*Ad hoc* list derived from PANAS-X (2 PE); Intensity	6 days
Baik, 2023	N= 30/63 (USA)	EMA; 8/day	*Ad hoc* list of adjectives (1 PE); Intensity	8 days

Abbreviations: DACL: Depression Adjectives Checklist; EMA: Ecological momentary assessment; GAD: generalized anxiety disorder; GDS: Geriatric Depression Scales; MDD: major depressive disorder; NA: Negative affect; NESDA-EMAA: The Netherlands Study of Depression and Anxiety – Ecological Momentary Assessment; PE: Positive Emotions; PANAS: Positive and Negative Affect Schedule; PANAS-X: Positive and Negative Affect Schedule extended form; USA: United States of America
[a] Samples are described by the number of participants with depression/total of participants and settings (country)

The ST-DACL is a self-reported questionnaire constituted by a list of 12 adjectives referring to "positive affects" and 22 adjectives referring to "negative affects" (Lubin, 1994). Of note, in the development paper and manual, no conceptual differences are mentioned between "affect", "feelings" and "emotions", they are used as synonyms for a transient trait. However, a difference is made with "mood", which is considered a state. Its development started in 1965 to provide a measure of "depressive mood" (trait) and "depressive feelings" (state) for both the general population and clinical population with mental disorders (Lubin, 1994; Lubin, 1965). The authors initially listed 171 adjectives and selected those that allow for the separation of two groups: people with and without depression. The state form asks participants how they feel today (Lubin, 1994). The trait form (DACL) asks participants how they feel generally (Lubin, 1994). Regarding the state form, the ST-DACL allows for the calculation of three scores: a state-positive score (summing the number of adjectives referring to "positive affects"), a state-negative score, and a mood-total score.

The PANAS is a self-report questionnaire consisting of 10 adjectives referring to "positive affects" and 10 adjectives referring to "negative affects", designed to describe mood states (Watson et al., 1988). The authors used the terms "affect" and "emotion" interchangeably, without clearly distinguishing between them. Its development was based on the two-level model of Watson and Tellegen in which the higher level reflects the valence of the mood (negative or positive states), while the lower level reflects the specific qualities of individual affects (Watson and Clark, 1999; Watson and Tellegen, 1985). Item reduction was performed using principal component analysis on 60 adjectives, retaining those that best discriminated between positive and negative valence (Watson et al., 1988; Zevon and Tellegen, 1982). Each affect is rated on a 5-point Likert scale ranging from "very slightly or not at all" to "very much". A 4-point Likert frequency rating was later shown to yield an equivalent factorial structure (Watson, 1988). The PANAS allows for the calculation of a Positive Affect (PA) and a Negative Affect (NA) score, each ranging from 10 to 50. Seven time frames are possible: moment, today, past few days, week, past few weeks, year, and in general (how you generally feel, how you feel on average), with the latter measuring a trait form (Watson, 1988).

Several versions of the PANAS were subsequently developed: a short form with 5 "positive" and 5 "negative affects" (PANAS-SF), an extended version of 60 items (PANAS-X), and a child version (PANAS-C) (Watson and Clark, 1999; Laurent et al., 1999; Thompson, 2007). In our review, two studies used the PANAS as a whole instrument and seven studies used ad hoc subsets of items from the PANAS-X (Table 1). Four of them cited Bylsma et al. to justify their choice of subset, although no rationale was provided in that paper (de Vos et al., 2017; Panaite et al., 2020; Bylsma et al., 2011; Bouwmans et al., 2017; Stavrakakis et al., 2015). Across studies based on PANAS items," positive emotions" were assessed using Likert scales of varying lengths, ranging from 5 to 100 points. (Table 1).

Finally, five studies used *ad hoc* lists of adjectives referred to as "positive emotions" the authors specifically developed for the study. All had either scarce or missing descriptions on how these lists were

developed. In particular, the NESDA-EMAA study used a list reported in a PhD thesis without precision on how it was developed (Bennik, 2015). Another study referred to the theoretical circumplex model of affect of Russel, which also inspired Watson et al when they developed the PANAS (Russell, 1980; Russell and Barrett, 1999; Shin et al., 2022). Three studies did not explain at all how their lists of "positive emotions" were developed (Baik and Newman, 2023; Minaeva et al., 2021; Peeters et al., 2006).

3.3. Type of "positive emotions" assessed across the studies

Across the 16 studies, 37 unique adjectives were used to measure what the authors described as "positive emotions" (Figure 2). However, several of these adjectives may not strictly represent emotions - understood here as valenced, discrete, short-lived states- but could instead be interpreted as affect (higher order dimensional tone) or mood (a longer-lasting context) depending on usage (e.g., euphoric, full of life, peaceful). Some adjectives might also be considered as cognitive (e.g., alert, attentive, interested), social (e.g., close to others, belonging), or behavioral descriptors (e.g., active, strong, vigorous). Figure 2 illustrates this heterogeneity by proposing a classification that reflects these alternative interpretations. Nevertheless, our proposed categorization should be regarded as provisional, since certain items could plausibly be represented by more than one category depending on theoretical or contextual considerations. What the authors described as "negative emotions" are reported for each study in Supplementary material 3 Table B. On average, the studies measured 7 "positive emotions", ranging from 1 to 14 with a median number of 6.5. The most measured emotions were "enthusiastic" and "happy", measured by nine studies, of which four measured both emotions. This was followed by "cheerful" (7 studies) and "satisfied" (6 studies). Eight "emotions" were measured by only one study: "close to others", "contented", "euphoric", "full of life", "hopeful", "in good spirits", "like you belong", and "self-assured".

3.4. COSMIN assessment of the development of the ST-DACL, the PANAS, the PANAS-X, and the PANAS-SF

Two standardized measurement instruments were identified by the systematic review: the ST-DACL and the PANAS. Both are patient-reported outcome measures (PROMs) and hence eligible for a COSMIN evaluation of their development. We also evaluated the PANAS-X (extended form) and the PANAS-SF (short form), since some studies reported having measured some of their adjectives. A detailed description of these tools is available in Table 1 and Supplementary material 4.

We included 17 publications to evaluate the development of the PROMs (Supplementary material 5). Table 2 and Table 3 report the results of the COSMIN evaluation regarding the quality of the development and the content validity respectively. While the ST-DACL was developed in samples of people with depression, this is not the case for the different form of the PANAS which were developed in samples of students.

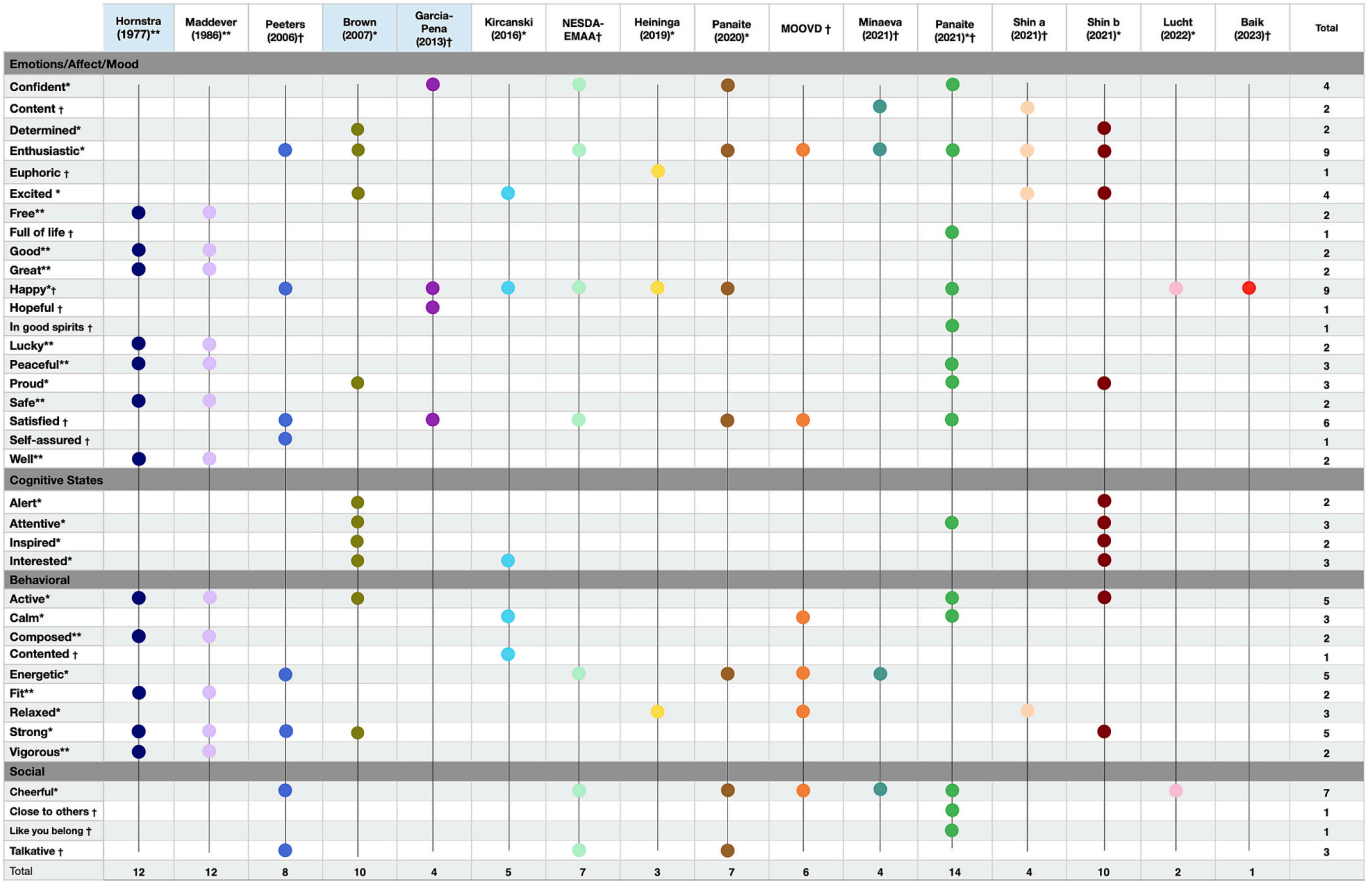

Fig. 2. Adjectives used to measure positive emotions across the 16 longitudinal studies of adults with depression included in the review. Numbers indicate how many studies assessed each adjective. Studies tagged with * used the full PANAS or a subset of items from the PANAS-X; adjectives tagged with * were derived from the PANAS or PANAS-X. Studies tagged with ** used the full ST-DACL; adjectives tagged with ** belong to the ST-DACL. Studies tagged with † used other scales (Kessler Psychological Distress Scale, K-10 (Kessler et al., 2002), or the Center for Epidemiologic Studies Depression Scale, CES-D) or ad hoc sets of adjectives; adjectives tagged with † were derived from the K-10, the CES-D, or were added ad hoc by the authors.
In Panaite et al. (2021), 'calm and peaceful' was reported as a single emotion by the authors, but we categorized it as two separate emotions, as we conducted a word-based qualitative analysis.

Regarding the quality of the development of the PROMS, the general design requirements (items 1 to 5) of the four PROMs were rated as "very good" for all items, except for the PANAS-X, which was found to have only an adequate description of its context of use (item 4). However, the concept elicitation requirements (items 6 to 13) were rated as "doubtful" for all PROMs. This was due to no comprehensive report of the methods used for concept elicitation being available in any of the retrieved publications. Thus, the quality of the design for the four PROMs was rated as "doubtful", in accordance with the COSMIN methodology, which requires the overall rating to be the lowest score of the features assessed. Regarding the quality evaluation of the cognitive interviews or equivalent method used to develop the PROMs, only the PANAS-SF reported to have used a focus group with participants and was thus rated as "very good" for including a sample representing the target population in its development (item 15). The other three measurement instruments were rated as "doubtful" as we found no report of the methods used in any of the retrieved publications. For all PROMs, there was no clear description, or too scarce information about the assessment of the comprehensibility (items 17 to 25) and comprehensiveness (items 26 to 35). Overall, reporting on scale development was incomprehensive, and the quality of the PROM development studies was therefore rated as "doubtful" for all four PROMs.

Regarding content validity, no study was specifically dedicated to this property. In particular, within the population of interest, no study involving people with depression examined the relevance, comprehensiveness, or comprehensibility of the items. The ST-DACL was the only scale to report surveys assessing comprehensibility across different educational levels, but this was conducted in the general population only (Lubin et al., 1990). Consequently, the quality of the PROM development studies was rated as 'doubtful' for all four instruments, due to: (1) the absence of a dedicated content validity study, (2) non-involvement of people with depression, and (3) the limited information on content validity provided in papers addressing other measurement properties. Given the poor ratings for development and content validity studies, there was insufficient evidence to conduct a meaningful assessment of content validity in samples of people with depression; therefore, we did not proceed, in line with COSMIN guidelines.

4. Discussion

This systematic scoping review included 16 longitudinal observational studies that reported to have monitored "positive emotions" in the course of adult depression. Of these, 81% used Ecological Momentary Assessment with heterogeneous collection modes, measurement instruments, types of "positive emotions" assessed, repetition of measures, duration of study, and incentives. Such heterogeneity in the procedures and measures of positive emotions does not facilitate the evidence synthesis of these longitudinal observational studies, which would be of importance to understanding the evolution of the disorder.

All studies reported having used lists of adjectives which could be rated either in terms of presence/absence or with Likert scales rating either intensity or frequency. These lists of adjectives corresponded either to standardized measures (ST-DACL, PANAS) that were developed to measure the mood rather than emotions, or to subsets of adjectives derived from the PANAS-X, or to *ad hoc* lists of adjectives assembled by the relevant study authors. Reporting on how adjectives were selected and whether and how global scores were calculated was often incomplete.

In total, 37 unique adjectives were used across the studies, of which a number are not strictly emotional. This overlap arises because a singular adjective (e.g., *peaceful*, *euphoric*, *full of life*) can be used to describe experiences of transient emotional episodes in certain circumstances, yet broader affective tones or mood states in others. Their interpretation depends notably on the situational and temporal frame in which they are assessed. Moreover, some adjectives might also be interpreted as cognitive, social or behavioral descriptors. In real life these dimensions

Table 2
Assessment of the development of the 4 Patient-Reported Outcome Measures (PROMS) retrieved by the search

	Design of the measures							Cognitive interview (CI) study				Total PROM development
	General design requirements							General design requirements	Comprehensibility	Comprehensiveness	Total CI study	
PROM	Clear construct	Clear origin of construct	Clear target population for which the measure was developed	Clear context of use	PROM developed in sample representing the target population*	Concept elicitation	Total design	CI study performed in sample representing the target population				
ST-DACL	V	V	V	V	V	D*	D*	D*	D*	D*	D*	D*
PANAS	V	V	V	V	V	D*	D*	D*	D*	D*	D*	D*
PANAS-X	V	V	V	A	V	D*	D*	D*	D*	D*	D*	D*
PANAS-SF	V	V	V	V	V	D*	D*	V	D*	D*	D*	D*

This table is based on the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) methodology for assessing the content validity of PROMs, user manual V.1.0. (https://www.cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf).
V, very good; A, adequate; D, doubtful; I, inadequate; NA, not applicable. Items marked with (*) were rated as "doubtful" due to insufficient reporting of concept elicitation or cognitive interview methods, not because of a mismatch between the construct and the measurement.

Table 3
Assessment of the content validity of the 4 Patient-Reported Outcome Measures (PROMS)

PROM	Content validity				
	Asking patients			Asking experts	
	Relevance	Comprehensiveness	Comprehensibility	Relevance	Comprehensiveness
ST-DACL	D	D	A	D*	D*
PANAS	D*	D*	D*	D*	D*
PANAS-X	D*	D*	D*	D*	D*
PANAS SF	D*	D*	D*	D*	D*

This table is based on the Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) methodology for assessing the content validity of PROMs, user manual V.1.0.
(https://www.cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf).
V, very good; A, adequate; D*, doubtful because of lack of study or lack of reporting; I, inadequate; NA, not applicable.

are often deeply intertwined: emotions are expressed through behaviors, cognitions can elicit emotions and vice versa, and social states are both shaped by and give shape to emotional experiences. Importantly, both the PANAS and ST-DACL were developed to assess mood or affect, not discrete emotions. There is limited evidence on the development of these PROMs and their content validity in samples of people with depression. This is why the quality of the PROM development study and content validity studies was rated as doubtful. Content-related concerns have already been discussed by the authors of these tools, such as the inclusion of non-emotional terms and the exclusion of obvious ones such as *happy* or *sad* (Watson and Clark, 1997). The authors of the PANAS responded by emphasizing that the scales were designed to measure positive and negative affect—higher-order dimensions of mood—rather than discrete emotions, and showed that adding such further emotional terms only modestly improved convergent validity (Watson and Clark, 1997). This underscores why cherry-picking PANAS items to represent "emotions" is methodologically problematic.

To our knowledge, this is the first systematic scoping review focusing on the procedures and measures used to assess positive emotions in longitudinal observational studies of adults with depression. Interestingly, while hand-searching for measurement instruments of positive emotions to prepare this work, we identified some alternatives, such as the Positive Emotions Scale (Shiota et al., 2006), which exclusively measures positive emotions (joy, contentment, pride, love, compassion, amusement, and awe). We also identified measurement instruments measuring both positive and negative affects, such as the Observed Emotion Rating Scale (Lawton et al., 1999); the Apparent Emotion Rating Instrument (Snyder et al., 1998) that measures six emotions: anger, anxiety, sadness, pleasure, interest, and tranquility; and the Modified Differential Emotion Scale (Galanakis et al., 2016). None of these measurement instruments were used in our sample of 16 studies, which instead relied on measurement instruments initially developed to measure moods (ST-DACL and PANAS).

This study has a number of limitations. A first limitation is the lack of dedicated search terms within bibliographic databases and the absence of validated search strategies for identifying the studies of interest. We addressed this problem by developing a list of relevant keywords validated by a professional librarian. A second limitation is that this study did not include interventional studies evaluating the evolution of emotions under treatment. This was a methodological choice since our group had already conducted a systematic search of clinical trial registries from 2018 to 2022 (Veal et al., 2024). This search retrieved 450 trials, of which 1.3% measured "positive emotions" either with the PANAS (5 trials) or the Profile of Mood States (POMS) (1 trial). A third limitation is that compliance rates and potential reactivity to repeated sampling were rarely reported, preventing us from benchmarking these aspects across studies. A fourth limitation concerns bipolar samples: they were sparse and heterogeneous, and factors such as manic symptoms or antidepressant exposure may influence positive affect. A final limitation is related to the use of the COSMIN guidelines to evaluate the development of the PROMS and its "loser takes all" decision rule which might understate the scales.

Regarding the research perspectives opened by our study, there is room to improve the knowledge about the longitudinal evolution of positive emotions in depressed patients in longitudinal observational studies. We tried to highlight the importance of the topic and the variety of methodological possibilities in terms of data collection modes, measurement instruments, repetition of measures and duration. With growing interest in EMA and the use of wearable devices, there is a need to ensure that individual studies can be readily included in evidence syntheses, which requires some level of harmonization and consistent reporting. This could start with the development of reporting guidelines for EMA studies.

With growing interest in EMA and the use of wearable devices, there is a need to ensure that individual studies can be readily included in evidence syntheses, which requires harmonization and consistent reporting. This could begin with the development of dedicated reporting guidelines for EMA studies. To our knowledge, no EMA reporting guideline has yet been developed using the highest methodological standards, besides two pioneering initiatives derived from systematic reviews of EMA for behavioral changes (Liao et al., 2016; Dao et al., 2021). As a preliminary step, Box 1 presents *"Minimum reporting items for EMA of emotional states—proposed for consideration."* We stress that Box 1 does not formally suggest a new reporting guideline. This would require a dedicated methodology, such as the Delphi consensus, in line with EQUATOR recommendations (https://www.equator-network.org/toolkits/developing-a-reporting-guideline/developing-your-reporting-guideline/, n.d.). Rather, it presents a pragmatic synthesis of key items. This Box should be viewed only as a preliminary contribution to stimulate debate, with the recognition that the development of robust reporting guidelines will require dedicated methodology.

Also, our study shows that there is no consensus on how to measure positive emotions. The use of adjectives drawn from scales designed to measure mood, with no explanation of how they were chosen, perpetuates the conceptual confusion between mood, feelings, emotions, and affects. The absence of clear measures for positive emotions reflects the inconsistent definition of the construct (Bringmann and Eronen, 2016). Hence, there is no consensus about the constituent features of positive emotions and their potential delineation from affects to define measurement within a "substantive formal theory" (Borgstede and Eggert, 2023). Furthermore, there is a need for clarification of the temporal resolution of the measurement of emotion: what time frame is needed for a specific purpose, such as diagnosis or clinical prediction, remains uncertain. For instance, if positive emotions are considered as part of an "emerging perceptual categorization" process (Barrett et al., 2007), then the question is what is the most useful temporal resolution for the measure (days, hours, minutes, seconds) and what type of indicator should be used (intensity, frequency, count). Of note, different types of temporal resolutions and indicators could carry different types of information in the context of dynamical processes (Kuppens et al., 2010), but can still be needed to address different purposes.

Regarding content validity, it is important to justify the use of each

Box 1
Presents preliminary reporting items to improve transparency and comparability of Ecological Momentary Assessment studies on emotional states. **For each item, researchers should both (i) describe what was done and (ii) justify their methodological choices.** It is not a formal reporting guideline, but reflects both the key fields we extracted and the information often missing or difficult to interpret in published studies.

1. Good practice

- **Preregistration and protocol availability**.
- **Patient and public involvement (PPI)**: describe whether and how lived-experience experts contributed to design and/or interpretation.
- **Open data and code**

2. Outcome

- **Domain**: specify what is measured (e.g., discrete emotions, affect, mood, behavioral manifestations of emotions, psychometric construct).
- **Measurement instrument**: name the tool (validated instrument, *ad hoc* item set, checklist) and provide exact wording of items. If *ad hoc*, describe the development process.
- **Time frame anchors**: wording used (e.g., "right now," "today," "during the past X hours").
- **Response scales**: scale type (Likert, VAS, slider), number of points, and whether intensity or frequency is assessed.
- **Aggregation methods**: describe how scores were calculated (e.g., item averages, summed scores, composites).
- **Contextual variables**: whether situational data were collected (e.g., location, activity, social context).
- **Incentives**: type (e.g., monetary, course credit, personalized feedback) and conditions.

3. Schedule / Sampling

- **Sampling plan**: fixed, random, semi-random, or event-based (e.g., when an emotion arises).
- **Prompt frequency**: number of prompts per day and minimum/maximum interval between prompts.
- **Clock time coverage**: start and end of assessment window during the day (e.g., 8:00–22:00).
- **Latency**: time allowed between prompt and response.
- **Duration**: e.g., total number of days assessed; number of waves if applicable.
- **Burden assessment**: whether participant burden or acceptability was assessed/reported.

4. EMA technology

- **Device/app**: medium of data collection (e.g., smartphone app, electronic diary, wearable device). Indicate any restrictions (e.g., requiring a personal smartphone) and whether time stamps were verified.
- **Prompting method**: type of reminder (e.g., SMS, push notification, alarm).
- **Participant preparation**: training, onboarding, or instructions provided.

5. Completeness

- **Adherence metrics**: proportion of prompts completed, compliance thresholds for inclusion in the analysis, dropout and attrition rates.
- **Missing data handling**: description of imputation or analytic strategy.

word for a scale. In a second step, after the development, it is necessary to evaluate the content validity as recommended by COSMIN regarding the relevance, comprehensibility, and comprehensiveness of the items included. Consensus about what positive emotions are and how to measure them should be found in order to compare and combine the results. Grounding the outcomes and their measures in patients' needs and experiences is a way of ensuring the clinical relevance of the findings. We should highlight efforts within the EMA/ESM research community to maintain an open searchable Item Repository that facilitate the inventory of existing measures (Kirtley et al., 2018).

With regard to the clinical perspectives of our study, we consider the monitoring of positive emotions as important as the monitoring of negative emotions. As clinicians, the therapeutic objective of managing depression aims not only to reduce the suffering caused by the omnipresence of negative emotions but also to restore positive emotions. This aligns with the definition of health by the World Health Organization (WHO), which emphasizes that health is not the absence of disease but a state of well-being, of which positive emotions are an important component. A diminution or absence of negative emotions does not imply the presence of positive emotions. However, while negative emotions are repeatedly monitored within depression scales, positive emotions rarely are assessed in studies and even less so in clinical practice.

Different clinical and research applications may require different EMA resolutions. For example, prognosis and relapse monitoring might rely on daily summaries of positive emotions, whereas mechanism-focused research (e.g., reward responsivity) may benefit from momentary assessments of intensity or variability. In treatment monitoring, indices such as inertia and instability may be more informative than mean levels, as they capture the dynamic patterns of emotional change that can signal treatment response or vulnerability to relapse. In addition, because sleep and circadian rhythms strongly influence daily emotional dynamics, integrating EMA protocols with sleep measures could provide valuable contextual information to interpret fluctuations in positive emotions (Socci et al., 2024).

More data are needed to understand the natural course of emotions in depression and identify potential phenotypes. Regarding interventional studies, monitoring positive emotions alongside negative ones might be of interest for both developing pathophysiological and psychopathological knowledge and evaluating treatment efficacy.

However, to prevent research waste, these evaluations should rely on valid procedures and measures, which remain to be determined.

In conclusion, because positive emotions are an important part of the patient's recovery, one critical next step would be to further explore this part of the disorder in research which requires a preliminary step of consensus on the procedures and measurement that would facilitate evidence synthesis.

CRediT authorship contribution statement

S. Dupuis-Lesavre: Writing – original draft, Formal analysis, Data curation. **C. Mebazaa:** Writing – review & editing, Formal analysis, Data curation. **C. Veal:** Writing – review & editing, Visualization, Formal analysis. **A. Ribeiro:** Writing – review & editing, Methodology, Data curation. **I. Boutron:** Writing – review & editing, Methodology, Conceptualization. **K. Krause:** Writing – review & editing, Methodology, Conceptualization. **Y. Morvan:** Writing – review & editing, Supervision, Methodology, Conceptualization. **A. Chevance:** Writing – review & editing, Validation, Supervision, Software, Resources, Project administration, Methodology, Data curation, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author (Astrid Chevance) used ChatGPT (OpenAI) in order to assist with editing the English. After using this tool, the author reviewed and edited the content as needed and takes full responsibility for the content of the published article.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

Declaration of competing interest

The authors report no conflict of interest.

Acknowledgements

We thank Colin Sidre, librarian at the Université de Paris Cité.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jad.2025.120387.

Data availability

The manuscript and supplementary material display all important data. Further extracted data will be shared on request to the corresponding author.

References

APA. American Psychological Association, 2018. [cited 2024 Oct 15]. Dictionary of Psychology. Available from: https://dictionary.apa.org/emotion.
Baik, S.Y., Newman, M.G., 2023 Apr. The transdiagnostic use of worry and rumination to avoid negative emotional contrasts following negative events: a momentary assessment study. J. Anxiety Disord. 95, 102679.
Barrett, L.F., Mesquita, B., Ochsner, K.N., Gross, J.J., 2007. The experience of emotion. Annu. Rev. Psychol. 58, 373–403.
Bassett, S.M., Schuette, S.A., O'Dwyer, L.C., Moskowitz, J.T., 2019 Nov. Positive affect and medication adherence in chronic conditions: a systematic review. Health Psychol. 38 (11), 960–974.
Bennik, E., 2015. Every Dark Cloud Has a Colored Lining: The Relation Between Positive and Negative Affect and Reactivity to Positive and Negative Events. University of Groningen, [Groningen].
Borgstede, M., Eggert, F., 2023 Feb 1. Squaring the circle: From latent variables to theory-based measurement. Theory Psychol. 33 (1), 118–137.
Bosley, H.G., Soyster, P.D., Fisher, A.J., 2019 Dec 30. Affect dynamics as predictors of symptom severity and treatment response in mood and anxiety disorders: evidence for specificity. J. Pers. Oriented Res. 5 (2), 101–113.
Bouwmans, M.E.J., Bos, E.H., Hoenders, H.J.R., Oldehinkel, A.J., de Jonge, P., 2017 Jan. Sleep quality predicts positive and negative affect but not vice versa. An electronic diary study in depressed and healthy individuals. J. Affect. Disord. 1 (207), 260–267.
Bringmann, L.F., Eronen, M.I., 2016 Feb 1. Heating up the measurement debate: What psychologists can learn from the history of physics. Theory Psychol. 26 (1), 27–43.
Brown, T.A., 2007 May. Temporal course and structural relationships among dimensions of temperament and DSM-IV anxiety and mood disorder constructs. J. Abnorm. Psychol. 116 (2), 313–328.
Bylsma, L.M., Taylor-Clift, A., Rottenberg, J., 2011 Feb. Emotional reactivity to daily events in major and minor depression. J. Abnorm. Psychol. 120 (1), 155–167.
Charvet, C., Boutron, I., Morvan, Y., Berre, C., Touboul, S., Gaillard, R., et al., 2022. How to measure mental pain: a systematic review assessing measures of mental pain. Evid.-Based Ment. Health Jul 28 (25) ebmental-2021.
Chevance, A., Ravaud, P., Tomlinson, A., Le Berre, C., Teufer, B., Touboul, S., et al., 2020 Aug. Identifying outcomes for depression that matter to patients, informal caregivers, and health-care professionals: qualitative content analysis of a large international online survey. Lancet Psychiatry 7 (8), 692–702.
Chevance, A., Tomlinson, A., Ravaud, P., Touboul, S., Henshall, C., Tran, V.T., et al., 2022 Dec. Important adverse events to be evaluated in antidepressant trials and meta-analyses in depression: a large international preference study including patients and healthcare professionals. Evid Based Ment. Health 25 (e1), e41–e48.
Clark, L.A., Watson, D., 1991. Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. J. Abnorm. Psychol. 100 (3), 316–336.
Dao, K., De Cocker, K., Tong, H., Kocaballi, A., Chow, C., Laranjo, L., 2021 Nov. Smartphone-delivered ecological momentary interventions based on ecological momentary assessments to promote health behaviors: systematic review and adapted checklist for reporting ecological momentary assessment and intervention studies. JMIR Mhealth Uhealth 9 (11), e22890. https://doi.org/10.2196/22890.
Dronavalli, M., Thompson, S., 2015 Jun. A systematic review of measurement tools of health and well-being for evaluating community-based interventions. J. Epidemiol. Community Health 3, 69.
DuBois, C.M., Lopez, O.V., Beale, E.E., Healy, B.C., Boehm, J.K., Huffman, J.C., 2015 Sep. Relationships between positive psychological constructs and health outcomes in patients with cardiovascular disease: a systematic review. Int. J. Cardiol. 15 (195), 265–280.
Ekman, P., Friesen, W.V., Ellsworth, P., 1972. Emotion in the human face: Guidelines for research and an integration of findings, xii. Pergamon Press, Oxford, England, p. 191 (Emotion in the human face: Guidelines for research and an integration of findings).
Fredrickson, B.L., Cohn, M.A., 2008. Positive emotions. In: Handbook of Emotions, 3rd ed. The Guilford Press, New York, NY, US, pp. 777–796.
Galanakis, M., Stalikas, A., Pezirkianidis, C., Karakasidou, E., 2016 Jan. Reliability and validity of the Modified Differential Emotions Scale (mDES) in a Greek sample. Psychology 1 (07), 101–113.
García-Peña, C., Wagner, F.A., Sánchez-García, S., Espinel-Bermúdez, C., Juárez-Cedillo, T., Pérez-Zepeda, M., et al., 2013 Sep 25. Late-life depressive symptoms: prediction models of change. J. Affect. Disord. 150 (3), 886–894.
Goodwin, G., Price, J., De Bodinat, C., Laredo, J., 2017. Emotional blunting with antidepressant treatments: A survey among depressed patients. J. Affect. Disord. 221. Oct 15 [cited 2024 Oct 15]. Available from: https://pubmed.ncbi.nlm.nih.gov/28628765/.
Hamaker, E.L., Wichers, M., 2017. No time like the present: Discovering the hidden dynamics in intensive longitudinal data. Curr. Dir. Psychol. Sci. 26 (1), 10–15.
Heininga, V.E., Dejonckheere, E., Houben, M., Obbels, J., Sienaert, P., Leroy, B., et al., 2019 Feb 8. The dynamical signature of anhedonia in major depressive disorder: positive emotion dynamics, reactivity, and recovery. BMC Psychiatry 19 (1), 59.
Hornstra, R.K., Klassen, D., 1977. The course of depression. Compr. Psychiatry 18 (2), 119–125.
https://www.equator-network.org/toolkits/developing-a-reporting-guideline/developing-your-reporting-guideline/ (consulted 09/06/2025).
Kessler, R.C., Andrews, G., Colpe, L.J., Hiripi, E., Mroczek, D.K., Normand, S.L.T., et al., 2002. Kessler Psychological Distress Scale [Internet]. https://doi.org/10.1037/t08324-000 [cited 2024 Oct 15]. Available from:
Kircanski, K., Thompson, R.J., Sorenson, J., Sherdell, L., Gotlib, I.H., 2018 Nov. The everyday dynamics of rumination and worry: precipitant events and affective consequences. Cognit. Emot. 32 (7), 1424–1436.
Kirtley, O., Eisele, G., Kunkels, Y.K., Hiekkaranta, A.P., Heck, L.V., Pihlajamäki, M., et al., 2018 Dec 4. The Experience Sampling Method (ESM) Item Repository [cited 2024 Dec 26]; Available from: https://osf.io/kg376/.
Kuppens, P., Oravecz, Z., Tuerlinckx, F., 2010. Feelings change: accounting for individual differences in the temporal dynamics of affect. J. Pers. Soc. Psychol. 99 (6), 1042–1060.
Laurent, J., Catanzaro, S., Joiner, T., Rudolph, K., Potter, K., Lambert, S., et al., 1999 Sep. A measure of positive and negative affect for children: scale development and preliminary validation. Psychol. Assess. 1 (11), 326–338.
Lawton, M., Haitsma, K., Perkinson, M., Ruckdeschel, K., 1999 Jan. Observed Affect and Quality of Life in Dementia: Further Affirmations and Problems. J. Ment. Health Aging 1 (5), 69–81.
Liao, Y., Skelton, K., Dunton, G., Bruening, M., 2016 Jun. A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: an adapted STROBE checklist for reporting EMA studies (CREMAS). J. Med. Internet Res. 18 (6), e151. https://doi.org/10.2196/jmir.4954.

Linton, M.J., Dieppe, P., Medina-Lara, A., 2016 Jul 1. Review of 99 self-report measures for assessing well-being in adults: exploring dimensions of well-being and developments over time. BMJ Open 6 (7), e010641.
Lubin, B., 1965. Adjective checklists for measurement of depression. Arch. Gen. Psychiatry 12 (1), 57–62.
Lubin, B., 1994. Manual for the depression adjective check lists (3rd edition). Education and Industrial Texting Service, San Diego, California.
Lubin, B., Collins, J.F., Seever, M., Van Whitlock, R., Dennis, A.J., 1990. Relationships among readability, reliability, and validity in a self-report adjective check list. Psychol. Assess. 2 (3), 256–261.
Maddever, H.M., Calhoun, K.S., 1986. A comparison of the mood and activities of depressed, previously depressed and non-depressed women. Behav. Res. Ther. 24 (3), 351–356.
Minaeva, O., George, S.V., Kuranova, A., Jacobs, N., Thiery, E., Derom, C., et al., 2021 Oct 11. Overnight affective dynamics and sleep characteristics as predictors of depression and its development in women. Sleep 44 (10) zsab129.
Mokkink, B.L., Prinsen, A.C.C., Patrick, L.D., Alonso, J., Bouter, M.L., de Vet, C.W.H., et al., 2018. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) - User Manual (Version 1.0).
Moskowitz, J.T., Cheung, E.O., Freedman, M., Fernando, C., Zhang, M.W., Huffman, J.C., et al., 2021. Measuring positive emotion outcomes in positive psychology interventions: a literature review. Emot. Rev. 13 (1), 60–73.
Ong, A.D., Kim, S., Young, S., Steptoe, A., 2017 Oct. Positive affect and sleep: a systematic review. Sleep Med. Rev. 35, 21–32.
Panaite, V., Rottenberg, J., Bylsma, L.M., 2020 Sep. Daily affective dynamics predict depression symptom trajectories among adults with major and minor depression. Affect. Sci. 1 (3), 186–198.
Panaite, V., Devendorf, A.R., Kashdan, T.B., Rottenberg, J., 2021 Mar 1. Daily life positive events predict well-being among depressed adults 10 years later. Clin. Psychol. Sci. 9 (2), 222–235.
Peeters, F., Berkhof, J., Delespaul, P., Rottenberg, J., Nicolson, N.A., 2006 Aug. Diurnal mood variation in major depressive disorder. Emotion 6 (3), 383–391.
Russell, J.A., 1980. A circumplex model of affect. J. Pers. Soc. Psychol. 39 (6), 1161–1178.
Russell, J.A., Barrett, L.F., 1999 May. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. J. Pers. Soc. Psychol. 76 (5), 805–819.
Shiffman, S., Stone, A.A., Hufford, M.R., 2008. Ecological momentary assessment. Annu. Rev. Clin. Psychol. 4, 1–32.
Shin, K.E., Newman, M.G., Jacobson, N.C., 2022 Jan. Emotion network density is a potential clinical marker for anxiety and depression: Comparison of ecological momentary assessment and daily diary. Br. J. Clin. Psychol. 61 Suppl 1 (Suppl. 1), 31–50.
Shiota, M.N., Keltner, D., John, O.P., 2006 Apr. Positive emotion dispositions differentially associated with Big Five personality and attachment style. J. Posit. Psychol. 1 (2), 61–71.
Snyder, M., Ryden, M.B., Shaver, P., Wang, J.J., Savik, K., Gross, C.R., et al., 1998 Mar 16. The apparent emotion rating instrument: assessing affect in cognitively impaired elders. Clin. Gerontol. 18 (4), 17–29.
Snyder, C.R., Lopez, S.J., Edwards, L.M., Marques, S.C. (Eds.), 2021. The Oxford Handbook of Positive Psychology [Internet]. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199396511.001.0001 [cited 2023 Sep 1]. Available from:
Socci, V., Festucci, F., Barlattani, T., Salfi, F., D'Aurizio, G., Rossi, R., et al., 2024. Sleep quality and emotional reactivity in patients with borderline personality disorder. Front Sleep [Internet]. 3 (2024).
Stavrakakis, N., Booij, S.H., Roest, A.M., de Jonge, P., Oldehinkel, A.J., Bos, E.H., 2015 Dec. Temporal dynamics of physical activity and affect in depressed and nondepressed individuals. Health Psychol. 34S, 1268–1277.
Terwee, C., Prinsen, A.C.C., Chiarotto, A., de Vet, C.W.H., Bouter, M.L., Alonso, J., et al., 2017. COSMIN Standards and Criteria for Evaluating the Content Validity of Health-related Patient-Reported Outcome Measures.
Terwee, C.B., Prinsen, CaC, Chiarotto, A., Westerman, M.J., Patrick, D.L., Alonso, J., et al., 2018 May. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual. Life Res. 27 (5), 1159–1170.
Thompson, E.R., 2007 Mar 1. Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). J. Cross-Cult. Psychol. 38 (2), 227–242.
Tricco, A.C., Lillie, E., Zarin, W., O'Brien, K.K., Colquhoun, H., Levac, D., et al., 2018 Oct 2. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann. Intern. Med. 169 (7), 467–473.
Veal, C., Tomlinson, A., Cipriani, A., Bulteau, S., Henry, C., Müh, C., et al., 2024 Apr. Heterogeneity of outcome measures in depression trials and the relevance of the content of outcome measures to patients: a systematic review. Lancet Psychiatry 11 (4), 285–294.
de Vos, S., Wardenaar, K.J., Bos, E.H., Wit, E.C., Bouwmans, M.E.J., de Jonge, P., 2017. An investigation of emotion dynamics in major depressive disorder patients and healthy persons using sparse longitudinal networks. PLoS One 12 (6), e0178586.
Watson, D., 1988. The vicissitudes of mood measurement: effects of varying descriptors, time frames, and response formats on measures of positive and negative affect. J. Pers. Soc. Psychol. 55 (1), 128–141.
Watson, D., Clark, L.A., 1997 Apr 1. Measurement and Mismeasurement of Mood: Recurrent and Emergent issues. J. Pers. Assess. 68 (2), 267–296.
Watson, D., Clark, L., 1999 Jan 1. The PANAS-X: Manual for the positive and negative affect schedule-expanded form. Psychology Publications. https://www2.psychology.uiowa.edu/faculty/clark/panas-x.pdf.
Watson, D., Tellegen, A., 1985 Sep. Toward a consensual structure of mood. Psychol. Bull. 98 (2), 219–235.
Watson, D., Clark, L.A., Tellegen, A., 1988 Jun. Development and validation of brief measures of positive and negative affect: the PANAS scales. J. Pers. Soc. Psychol. 54 (6), 1063–1070.
Zevon, M.A., Tellegen, A., 1982. The structure of mood change: an idiographic/nomothetic analysis. J. Pers. Soc. Psychol. 43 (1), 111–122.