Published in final edited form as:
Brain Behav Immun. 2020 August ; 88: 887–900. doi:10.1016/j.bbi.2020.04.078.

The effect of acute stress on salivary markers of inflammation: A systematic review and meta-analysis

Yvette Z. Szabo[a,b,c], **Danica C. Slavish**[d], **Jennifer E. Graham-Engeland**[e]

[a]Department of Veterans Affairs VISN 17 Center of Excellence for Research on Returning War Veterans, Waco, TX, USA

[b]Texas A&M College of Medicine, Bryan, TX, USA

[c]Department of Health, Human Performance and Recreation, Baylor University, Waco, TX, USA

[d]Department of Psychology, University of North Texas, Denton, TX, USA

[e]Department of Biobehavioral Health, The Pennsylvania State University, University Park, PA, USA

Abstract

Background: Salivary biomarkers of inflammation are increasingly used in stress research. This systematic review and meta-analysis provides a quantitative summary of changes in salivary inflammatory markers in response to acute stress.

Method: The review included 1,558 participants (42 unique samples, 33 studies) obtained through electronic databases (PubMed, PsycINFO, Embase), reference treeing, and articles identified by a 2015 review on a similar topic. To be eligible, articles had to be quantitative and assess change in at least one biomarker of salivary inflammation in response to acute stress in adults. The primary outcome was magnitude of change in inflammatory biomarkers (Cohen's *d* for repeated measures [d_{av}]).

Results: Measures of salivary inflammation included: C-reactive protein (CRP), granulocyte-macrophage colony-stimulating factor (GM-CSF), interleukin (IL)-1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12p70, IL-13, IL-17A, IL-18, IL-21, interferon (IFN)-α, INF-γ, and tumor necrosis factor (TNF)-α. Cytokines IL-6 (k = 26, d_{av} = 0.27), IL-10 (k = 11, d_{av} = 0.34), TNF-α (k = 10, d_{av} = 0.57), and INF-γ (k = 6, d_{av} = 0.28) significantly increased in response to stress. Post hoc sensitivity analyses revealed that IL-1β (k = 19, d_{av} = 0.16) and IL-8 (k = 7, d_{av} = 0.30) also increased post-stress, but findings with INF-γ did not hold, after removing one outlier study each. Examination of moderators suggested that study methodology and sample demographics moderated some associations.

Corresponding Authors: Yvette Z. Szabo, 4800 Memorial Drive (151C), Waco, Texas 76711 (254) 297-3179; yvette.szabo@va.gov; Danica C. Slavish, 1155 Union Circle #311280, Denton, TX 76203; (608) 302-5682; danica.slavish@unt.edu.

Systematic review registration number: 125121

Disclosures
[All authors] reported no biomedical financial interests or potential conflicts of interest.

Conclusions: This meta-analysis revealed that certain inflammatory cytokines increase in response to acute stress. Significant heterogeneity in results and moderator analyses suggest need for standardization of research protocols. Directions for future research are discussed.

Keywords

saliva; inflammation; stress; TSST; meta-analysis; systematic review

1. Introduction

Inflammation has been recognized as an important correlate of mental and physical health (1). Inflammation is a non-specific biological cascade initiated to fight pathogens and promote healing. During tissue damage, microbial invasion, or exposure to foreign particles, inflammatory biomarkers such as cytokines and acute phase proteins are released from multiple cells, including macrophages, neutrophils, and endothelial cells (2). Although such inflammatory responses can be adaptive in the short term, chronic or excessive inflammation can lead to tissue damage, slowed wound healing, dampened ability to fight infection, and can put an individual at risk for chronic diseases (3–10). Laboratory-based studies in humans have demonstrated that inflammatory biomarkers are stress-responsive (11–13). Acute stressor paradigms can be implemented to assess acute changes in inflammatory biomarkers with biomarkers assessed before and after the stressor. Such stressor paradigms may include psychological stressors (e.g., a social evaluative speech or cognitive tasks) or exercise or physical stressors (e.g., the cold pressor test or exercise). The degree of psychological and immune reactivity in response to acute stress may reflect vulnerability to disease, as it has been shown to predict depression symptoms and susceptibility to infectious disease (14–16).

In controlled settings, the gold standard is to use blood samples to determine levels of inflammatory biomarkers. Systemic inflammation typically is measured by assessing levels of either pro-inflammatory cytokines (e.g., interleukin [IL]-1β, tumor necrosis factor [TNF]-α), anti-inflammatory cytokines (e.g., IL-10), or acute phase proteins (e.g., C-reactive protein [CRP]). However, there is an increasing interest in the ability to assess biological markers of stress reactivity in humans using less invasive methods.

Recent work has explored saliva as a medium in which to assess inflammatory changes in response to stress. Saliva consists of mucous and fluid secreted from the salivary glands, gingival fold, oral mucosa transudate, and mucous from the nasal cavity (17). Compared to blood, saliva is less invasive, cheaper, safer, and less burdensome to collect, making it a desirable medium in which to assess inflammatory biomarkers in and outside of laboratory settings (18). Although salivary markers of inflammation have been inconsistently associated with blood-based markers, they appear to have predictive utility for disease states. For example, salivary cytokines show cross-sectional associations with measures of mental health (e.g., depressive symptoms, post-traumatic stress symptoms, vital exhaustion) (19–21). Salivary markers like CRP and IL-6 also are associated with chronic diseases such as acute myocardial infarction (22), cardiovascular disease (23), metabolic syndrome (24), and rheumatoid arthritis (25).

A recent narrative review of 13 studies assessing salivary markers of inflammation in response to acute stress suggests that inflammatory biomarkers IL-6, TNF- α, and IL-1β appear to change in response to acute psychological and exercise laboratory stressors (13). In this review, Slavish and colleagues provided guidelines for future research and posed possible moderators of these responses. Since the 2015 review, the literature examining changes in salivary inflammation following acute stress has grown rapidly. The current research built on the 2015 review by conducting a systematic review and meta-analysis and answering four empirical questions: 1) Which salivary inflammatory markers reliably change following exposure to acute stress? 2) At what time point do inflammatory markers in saliva exhibit the largest changes from pre- to post-stressor? 3) What factors related to individual differences between participants influence patterns of salivary inflammatory responses to stress? 4) How do methodological factors influence these patterns of responses to stress? By answering these questions, we sought to summarize evidence, contextualize the current findings, inform methodology in the field, and enable recommendations for future research.

2. Method

The review was conducted using Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (26,27) with consultation of protocols for meta-analyses (i.e., Meta-Analysis of Observational Studies in Epidemiology [MOOSE]) (28). The review was registered at PROSPERO (International Prospective Register of Systematic Reviews; registration number: 125121/CRD42019125121]), and a protocol was published to prespecify the methods used in the systematic review and meta-analysis (29).

Inclusion criteria matched those in the published protocol (29). Briefly, each article had to be quantitative and assess change in at least one biomarker of salivary inflammation in response to an acute stressor in a sample of adults. Although we originally intended to consider it, we did not include secretory immunoglobulin A (sIgA), given that it appears to have only indirect linkages with inflammation (30).

First, the 13 articles included in the 2015 review were obtained. Then an electronic database search of PsycINFO, PubMed, and Embase was conducted using the following keywords: “acute stress OR stress* OR task OR challenge” AND “saliva*” AND “inflammat* OR interleukin OR cytokine OR fibrinogen OR C-reactive protein.” Minor changes were made to the sample search strategy reported in the protocol (29), specifically the use of quotations around terms and filters were not set to English (Supplemental Table 2). We also included the 117 studies citing the 2015 review, obtained through Google Scholar. Dates of coverage were through December 16, 2019 (original pre-registered date was March 31, 2019).

2.1 Study Selection:

Two authors (YZS and DCS) screened the titles, abstracts, and full articles for inclusion criteria using a two-stage process. All discrepancies were resolved by discussion, and any difficulty in consensus was referred to a third author (JEG). First, after training on a set of 50 abstracts, a set of four “yes/no” questions were applied to screen titles and abstracts: 1) report quantitative data, 2) were conducted in human adults, 3) used an acute stressor, and 4)

assessed at least one inflammatory biomarker in saliva in response to acute stress (97.1% agreement, $\kappa = 0.78$). One article was in Italian, and the author of this article helped determine it did not meet study criteria. Full-text articles ($n = 141$) then were obtained for articles meeting screening criteria and for the 13 articles included in the 2015 review. Articles that could not be obtained through the databases were attempted to be retrieved by contacting study authors; all but 1 were received or determined to be conference presentations (and thus ineligible). Then, inclusion and exclusion criteria were applied to the full-texts after training on a set of five articles (91.1% agreement, $\kappa = 0.81$). When multiple publications of the same sample were identified ($k = 5$), the larger sample or more recent publication was included.

2.2 Data Items:

Studies meeting inclusion criteria were independently coded. Data were extracted using a standardized Excel spreadsheet and codebook (29). In addition to the coding guide presented in the protocol paper, we added the following items: % White (pre-registered but not included in the original coding guide), % college educated, used kit designed for saliva (yes / no), body mass index, and additional biomarkers. Discrepancies were discussed and resolved by two authors (87.4% agreement); a third author was consulted to in the case of difficulty in consensus. Authors were contacted when necessary information was not reported.

2.3 Study Quality and Risk of Bias.

Risk of bias was assessed at the study level and outcome level. At the study level, two raters independently assessed quality and risk of bias for each study using a questionnaire containing nine items (64.8% agreement). This questionnaire was developed from the Risk Of Bias In Non-Randomized Studies - of Interventions (ROBINS-I) criteria and other published guides (see protocol paper for description) (29); slight modifications were made to the quality guide upon reviewing the studies (Supplemental Table 1). For data not reported, authors were contacted. For those who did not respond, the numerical rating of the published information was utilized. A narrative summary of study quality and risk of bias was conducted, and the total study quality score was used as a moderator of analyses. Study publication bias was evaluated using a contour-enhanced funnel plot to depict reporting bias of study results (31).

2.4 Data Analysis Plan

2.4.1 Synthesis of results:

An omnibus random effects meta-analysis examined changes in each inflammatory biomarker with at least four unique samples using Cohen's d for dependent samples (Cohen's d_{av}, see equation 10 in Lakens 2013) (32). Although we initially pre-registered to run meta-analyses on studies with as few as two unique samples, to reduce the chance of reporting unreliable findings we only interpret studies with four or more unique studies. Effect sizes were calculated so positive values represented increases from baseline to post-stress, whereas negative values represented decreases from baseline to post-stress. The variance of each effect size was calculated using equation 5 from Hirst et al., 2018 (33). In keeping with our pre-registered hypotheses, for inflammatory biomarkers assessed at multiple time points in the same study, we used the largest effect size in the

omnibus meta-analysis. This approach has three main advantages: 1) it reduces the chance of making a Type 1 error by including only one effect size per biomarker; 2) it reduces potential "wash-out" effects by excluding more distal "recovery" time points when salivary inflammation levels may have returned to baseline; and 3) it is consistent with our goal of determining biomarkers that can be reliably assessed and which show the largest magnitude of change. For studies with both positive and negative effect sizes, we chose the *largest positive* effect size. For studies with only negative effect sizes, we chose the *smallest negative* effect size. All effect sizes for all biomarkers and time points are reported in Table 1. Effects from studies with multiple within-person conditions ($k = 3$) were aggregated using the Borenstein, Hedges, Higgins and Rothstein (BHHR) procedure (34).

2.4.2 Additional analyses: Analyses examined psychosocial, demographic, and methodological moderators of salivary inflammatory responses to acute stress. Though not specified in the protocol paper, a decision was made before analyses were conducted to not conduct moderator analyses for biomarkers that had fewer than five unique samples (i.e., IL-5, IL-12p70, IL-13, IL-17a, interferon [INF]-α, and granulocyte-macrophage colony-stimulating factor [GM-CSF]), to reduce Type I error (12,35). Pre-registered demographic and psychosocial moderators were: gender/sex (% female), average sample age, race/ethnicity (% white), salivary flow rate (0 = not assessed, 1 = assessed), oral health status (0 = not assessed, 1 = assessed), and general health status (1 = Healthy; 2 = Clinical; 3 = Mixed) (20). Pre-registered methodological moderators included: type of stressor (1 = TSST; 2 = speech task other than TSST; 3 = exercise; 4 = Stroop/cognitive; 5 = cold pressor; 6 = other), sample timing (continuous number of minutes post-stressor that the maximum effect size was found), assay technique (1 = single plex ELISA or bead array; 2 = multiplex ELISA or bead array), and overall study quality (continuous score from 0 to 15) (29). Post-hoc analyses (i.e., not pre-registered) included running sensitivity analyses removing influential outlier cases (36) and examining stressor length (total number of minutes) as a moderator.

3. Results

3.1 Search

The database search returned 1,897 citations, and 102 records were identified through other sources. Removal of duplicates yielded 1,698 unique abstracts and titles. Of these, 141 articles were eligible for full-text screening. A total of 38 articles were eligible for synthesis but we did not receive relevant information on 5 studies; thus, 33 studies, with 42 unique samples and 1,558 participants, met full eligibility and were included in the meta-analysis. The PRISMA flow diagram of the screening process is shown in Supplementary Figure 1.

3.2 Study Description

Table 1 presents sample characteristics and study measures. Sample sizes ranged from 7 to 115 participants. A total of 17 analytes were studied. The most frequently measured biomarker was IL-6 ($k = 26$, $n = 991$), followed by IL-1β ($k = 20$, $n = 723$), IL-10 ($k = 11$, $n = 327$), TNF-α ($k = 10$, $n = 311$), IL-8 ($k = 8$, $n = 186$), CRP ($k = 6$, $n = 243$), IFN-γ ($k = 6$, $n = 139$), IL-2 ($k = 5$, $n = 87$), IL-4 ($k = 5$, $n = 87$), IL-5 ($k = 4$, $n = 78$), IL-12p70 ($k = 4$, $n = 121$). Each GM-CSF ($k = 2$, $n = 18$), IL-13 ($k = 2$, $n = 60$), IL-17A ($k = 2$, $n = 60$), and IFN-

α (k = 2, n = 60), and. Forest and funnel plots for these studies are presented in the supplemental materials (Supplemental Figures 3A–B, 12A–B, 13A–B and 14A–B). Both IL-18 (k = 1, n = 61) and IL-21 (k = 1, n = 20) were examined in one study and could not be meta-analyzed. Due to limited assessment of these makers in saliva, no definitive conclusions can be made about their stress reactivity. The average stressor length across all studies was 28.10 minutes (SD = 35.46, range = 2 to 180 minutes). Table 2 shows descriptive information on the timing of the peak effect size and average stressor length by biomarker. The average number of minutes after the start of the stressor that the peak effect size was observed ranged from 30 minutes to 100 minutes, depending on the specific biomarker.

3.3 Results of Individual Studies

Weighted effect sizes and 95% CIs for each sample and biomarker are shown in Supplemental Figures 2a through 16a. The overall average effect sizes for IL-6, IL-10, IFN-γ and TNF-α, were significant, positive, and small to moderate in magnitude, suggesting these inflammatory biomarkers increased modestly in response to acute stress. The heterogeneity statistic for IL-1β, IL-2, IL-4, IL-6, and TNF-α was statistically significant, suggesting heterogeneity across the samples was more than would be expected by chance alone.

Two cytokines, IL-18 and IL-21, were only examined in a single study. La Fratta and colleagues (19) measured IL-18 30 minutes before and 30 minutes after a half hour academic exam among 61 male students. Levels of IL-18 significantly decreased from pre- to post- exam (d_{av} = −.89). Moreira and colleagues (37) measured IL-21 in 20 male basketball players competing in a championship basketball game. Saliva samples were taken before a 30 minute warm up and 10–15 minutes after an approximately 120 minute basketball game. Levels of IL-21 significantly decreased from pre- to post-game (d not reported).

3.4 Risk of Bias and Study Quality

Publication bias was evaluated by inspecting contoured funnel plots (Supplemental Figures 2b–16b). Funnel plots were fairly symmetrical for most biomarkers, which suggests publication bias does not seem to be affecting the results. Of the markers with 5 or more studies, some (IL-1β, IL-2, IL-4, IL-6, IL-8, IL-10, and TNF-α) showed 1 or 2 studies outside the range of the funnel plot. Sensitivity analyses (below) were conducted to examine the potential effect of these outliers on results.

Total study quality scores ranged from 3–13 (see Table 1). The majority of the samples (95.23%) did not select participants representative of the general or target population. Comorbid health conditions were primarily either covaried in analyses (42.86%) or analyzed separately (40.48%). Just under half of the studies (42.86%) were powered to detect a small effect or reported an *a priori* power analysis. Just under half of the studies used non-validated stressors (47.62%). However, most studies (88.10%) used a separate self-report or physiological measure as a manipulation check. Additionally, 90.48% reported appropriate procedures for the storage and processing of samples. Only one study assayed samples in

triplicate, but just over half assayed sample in duplicate (57.14%), while 40.48% either assayed samples in singulate or did not report these data. No studies failed to report on a hypothesized outcome or incorrectly interpreted their findings, and 64.29% interpreted the outcome of interest well (i.e., commented on stress reactivity of the inflammatory biomarker in the discussion). Half of the studies (50.00%) did not report missing data, had high amounts of missing data, or did not report use of advanced techniques to handle missing data.

3.5 Additional Analyses

3.5.1 Moderator Analyses.—

Moderator analyses were conducted for studies that had at least 5 unique samples: CRP, IFN- γ, IL-1β, IL-2, IL-4, IL-6, IL-8, IL-10, and TNF-α. Significant moderator analyses are discussed below; additional moderation results can be found in Supplemental tables 3 through 13.

For IL-1β, samples with a lower percentage of participants who identified as white (i.e., samples with a higher percentage of racial or ethnic minority participants) yielded higher effect sizes for IL-1β: For every additional 1% of the sample that identified as white, effect sizes decreased by 0.05 units ($b = -0.05$, $SE = 0.02$, $p = .002$, $k = 13$). There was also a significant intercept, such that samples comprised of exclusively minority participants had an average effect size of 3.44 ($b = 3.44$, $SE = 1.11$, $p = .002$).

For IL-2, stressor type was a significant moderator, such that non-TSST speech tasks produced effect sizes 4.74 units greater compared to any other type of stressor ($b = 4.74$, $SE = 1.23$, $p < .001$, $k = 5$). Flow rate assessment was also a significant moderator: Studies that assessed flow rate produced effect sizes that were 4.87 units larger ($b = 4.87$, $SE = 0.98$, $p < .001$, $k = 5$). Oral health assessment was a significant moderator: Studies that assessed oral health produced effect sizes that were 4.87 units larger ($b = 4.87$, $SE = 0.98$, $p < .001$, $k = 5$). Of note, the only study that used a non-TSST speech task, assessed flow rate, and assessed oral health was the influential study (i.e., outlier) for IL-2. Sample timing also was a statistically significant moderator: For every one-minute later after the start of the stressor the peak IL-2 effect was observed, effect sizes increased by 0.03 units ($b = 0.03$, $SE = 0.01$, $p < .001$, $k = 5$). Finally, stressor length was a significant moderator for IL-2; each additional minute in stressor length yielded effect sizes 0.04 units larger ($b = 0.04$, $SE = 0.02$, $p = .01$, $k = 5$).

For IL-4, stressor type was a significant moderator, with non-TSST speech tasks producing effect sizes 3.95 units larger compared to any other type of stressor ($b = 3.95$, $SE = 0.88$, $p < .001$, $k = 5$). Flow rate assessment was a significant moderator: Studies that assessed flow rate produced effect sizes that were 4.15 units larger ($b = 4.15$, $SE = 0.86$, $p < .001$, $k = 5$). Oral health assessment was a significant moderator, such that studies that assessed oral health produced effect sizes that were 4.15 units larger ($b = 4.15$, $SE = 0.86$, $p < .001$, $k = 5$). Of note, the only study that used a non-TSST speech task, assessed flow rate, and assessed oral health was the influential study (i.e., outlier) for IL-4. Sample timing also was a statistically significant moderator: For every one-minute later after the start of the stressor the peak IL-4 effect was observed, effect sizes increased by 0.02 units ($b = 0.02$, $SE = 0.01$, $p < .001$, $k = 5$). Finally, stressor length was also a significant moderator for IL-4: Each

additional minute in stressor length yielded effect sizes 0.04 units larger ($b = 0.04$, $SE = 0.02$, $p = .02$, $k = 5$).

For IL-8, assay type was a significant moderator. Studies that used a multiplex ELISA or bead array resulted in effect sizes that were 1.41 units greater ($b = 1.41$, $SE = 0.49$, $p < .01$, $k = 8$). Sample timing was also a statistically significant moderator for IL-8: For every one-minute later after the start of the stressor the peak IL-8 effect was observed, effect sizes decreased by 0.01 units ($b = -0.01$, $SE = 0.002$, $p = .03$, $k = 6$).

For TNF-α, compared to any other type of stressor, speech tasks other than the TSST yielded effect sizes that were 2.74 units greater ($b = 2.74$, $SE = .94$, $p < .01$, $k = 10$). Sample timing also was a statistically significant moderator: For every one-minute after the start of the stressor the peak TNF-α effect was observed, effect sizes increased by 0.01 units ($b = 0.01$, $SE = .003$, $p < .001$, $k = 8$). Finally, stressor length was also a significant moderator TNF-α: Each additional minute in stressor length yielded effect sizes 0.03 units larger ($b = 0.03$, $SE = 0.01$, $p < .001$, $k = 8$).

For CRP, IFN-γ, IL-6, and IL-10, there were no significant moderators. In addition, moderators with percent of the sample that identified as female could not be fully tested due to missing data (IL-10, $k = 3$; TNF-α, $k = 3$; IL-2, $k = 3$, IL-4, $k = 3$, IFN-γ, $k = 4$). Additionally, moderator analyses could not be conducted for IFN-γ for % white, sample timing, or sample length ($k = 4$ for each of these analyses). Further, many moderators could not be assessed due to all included studies not measuring oral health or flow rate (oral health: IL-2, IL-4, IL-8; flow rate: IL-2, IL-4) and all included studies using multiplex assays (IL-2, IL-4, IL-8, TNF- α).

3.5.2 Sensitivity Analyses.—

Sensitivity analyses were conducted using the "influence()" function in the R package *metafor* (36) for studies that had at least 5 unique samples: CRP, IFN- γ, IL-1β, IL-2, IL-4, IL-6, IL-8, IL-10, and TNF-α. Studies considered to be influential outliers (36) were removed from analyses (IFN-γ [$k = 1$], IL-1β [$k = 1$], IL-2 [$k = 1$], IL-4 [$k = 1$], IL-6 [$k = 1$], IL-8 [$k = 1$], and TNF-α [$k = 2$].

Removing these outliers did not change the pattern or significance of findings for the stress reactivity of IL-6 ($d_{av} = 0.22$, $SE = 0.05$, $p < .001$), TNF-α ($d_{av} = .26$, $SE = .08$, $p = .002$), IL-2 ($d_{av} = .20$, $SE = .16$, $p = .21$), and IL-4 ($d_{av} = .06$, $SE = .16$, $p = .71$). However, removing outliers did change the pattern of results for IL-1β and INF-γ. Specifically, the original non-significant effects of acute stress on IL-1β and IL-8 became significant (IL-1β: $d_{av} = .16$, $SE = .08$, $p < .05$; IL-8: $d_{av} = .30$, $SE = .11$, $p = .01$), and the original significant effect of INF-γ became non-significant ($d_{av} = .18$, $SE = .14$, $p = .19$). There were no influential studies for CRP or IL-10, so sensitivity analyses were not conducted.

4. Discussion

This systematic review and meta-analysis observed a small to moderate increase in salivary IL-6, IL-10, IFN-γ and TNF-α from pre- to post-stress, but no significant changes in salivary CRP, IL-1β, IL-2, IL-4, IL-5, IL-8, IL-13, IL-17A, IFN-α or GM-CSF. For TNF-α and IL-6, our findings are consistent with Slavish and colleagues (13). We expanded on this

previous review by providing data and reporting significant effects for IL-10 and IFN-γ. Our findings with CRP are consistent with previous meta-analyses in blood (11,12) and provide strong evidence that CRP is not reliably reactive to acute stress in saliva. Our findings with IL-1β differ from previous syntheses in saliva (13) and blood (11,12). However, our sensitivity analyses suggest that our lack of significant finding with IL-1β may be driven by one outlying study. When this study was removed post hoc, IL-1β demonstrated a moderate and highly significant increase from pre- to post-stress across the remaining studies. Further, although Marsland and colleagues noted possible publication bias for IL-1β (such that significant results appeared more likely to have been published than null effects) in their 2017 review using blood-based inflammatory markers, we did not detect evidence of publication bias for IL-1β in the present review after removing the one outlying study. In addition, our sensitivity analyses suggested that the significant findings we observed with INF-γ did not hold and that IL-8 demonstrated significant increases after removing outlying studies. Overall, our analyses demonstrate that several inflammatory biomarkers increase in response to acute stress and may serve as valuable markers of stress reactivity. Further, some biomarkers were also influenced by sample demographics and study methodology, as interpreted below.

4.1 Summary of Evidence

Using Grading of Recommendations Assessment, Development and Evaluation (GRADE) recommendations (38), the quality of the studies was moderate to high. Further research may be likely to have an impact on our confidence in the estimate of our effects, particularly for the biomarkers with fewer than 10 studies included (35). For risk of bias, we have moderate confidence in our results, and future research is only somewhat likely to change the confidence in the estimates of our effects. Examination of funnel plots and sensitivity analyses suggest that publication bias likely did not contribute to the observed effect sizes and that study characteristics contributed to heterogeneity observed. However, almost no studies randomly sampled from the general or target populations, suggesting results may lack generalizability. A strength is that most studies used a separate measure as a manipulation check (e.g., cortisol, heart rate blood pressure, subjective stress reports). However, many articles did not report missing data or quantity of samples with biomarker levels below detectable limits, a common issue when examining salivary markers of inflammation (39–41). As with many markers of inflammation, it can be difficult to obtain detectable levels in relatively healthy samples. Values below detectable limits are data missing not at random (MNAR), which poses a unique challenge for data analysis. Multiple imputation techniques are not recommended for MNAR (42). Therefore, when inflammatory data are missing due to levels being too low to detect, researchers may want to consider imputing a very small value for missing data (i.e., the value of half the lower limit of detection). This is a common strategy when examining salivary markers of inflammation (39–41,43) that capitalizes on using all available information.

Despite heterogeneity in effect sizes from study to study, the significant effects of acute stress on salivary IL-6, IL-10, and TNF-α, appear to be robust and unaffected by outliers. Similarly, for CRP the quality of evidence is high: Thus, future research is unlikely to change the emerging consensus that salivary CRP appears not to increase reliably in

response to acute stress. Findings for INF-γ, however, should be considered preliminary, as these analyses did not remain statistically significant after sensitivity analyses. For IL-1β, there is moderate confidence that IL-1β may increase in response to acute stress; we initially observed a null finding with IL-1β that became statistically significant when removing an outlier, and the observation of a significant increase in IL-1β in response to stress is consistent with previous research (11–13). Findings for IL-8 should also be considered preliminary, as we initially observed a null finding with IL-8 that became statistically significant when removing an outlier.

4.2 Moderator Analyses

Demographic moderators examined yielded only a few, albeit potentially important, findings. For IL-1β, the percentage of the sample that identified as white was negatively associated with stress-related increases. This suggests that samples with a great proportion of racial or ethnic minorities exhibited larger effect sizes for IL-1β. Previous research has identified racial or ethnic differences in stress-reactivity of serum IL-6 (38). This is a novel finding for IL-1β, which suggests that racial or ethnic identity will be important for future research to recruit and consider. It will also be important for researchers to examine the predictive validity of what these potential racial or ethnic differences in stress-reactivity may mean for health outcomes. Contrary to our hypotheses, other demographic factors – such as age, health status, and the percentage of the sample that identified as female – were not significant moderators of any biomarker assessed. Our ability to detect moderation by age may be due to a restricted range included in this sample (i.e., most participants were young adults). Further, few studies used clinical samples, and samples with diverse mental and physical health conditions were recruited. Thus, more research focused on testing age and clinical health status may be helpful to inform the role of these factors in stress-related changes in salivary inflammation. Although no evidence emerged for differences in stress reactivity based on gender, few studies tested this directly. Moreover, the role of the menstrual cycle, which can affect stress reactivity (44), was not assessed in this review and may be important in understanding sex/gender differences in stress-related increases of salivary inflammatory markers. Finally, flow rate and oral health emerged as significant moderators only for IL-2 and IL-4. However, many studies did not report this information, and therefore, definitive conclusions about the role of these factors cannot be drawn.

In terms of methodological moderators, speech tasks that were not the TSST resulted in larger effect sizes for TNF-α, IL-2, and IL-4, and exercise stressors were associated with smaller increases in IL-10 (compared to any other type of stressor). However, this review was limited by the large number of stressors used and the variability of validation of these stressors in previous research. As additional studies utilize salivary measures and diverse stressors, new patterns may be observed and new hypotheses can be tested. As an example, two studies (45,46) used variations of cold pressor tasks; one of these studies employed a traditional cold pressor task and the other study used a “socially evaluative” cold pressor task. Using our preregistered coding guide, the traditional cold pressor task was coded in the cold pressor task category, while the socially evaluative cold pressor task was categorized as an “other” type of stressor (due to its combination of physiological and psychosocial elements). Interestingly, as shown in Table 1, IL-1β decreased in response to both cold

pressor-related stressors, which may be due to an anti-inflammatory impact of the cold water. With more research, new patterns in the most “potent” type of stressor for eliciting changes in salivary markers of inflammation may emerge.

Our results indicated that timing of sample collection was a significant moderator for many studies, but findings were somewhat inconsistent: A longer delay between the start of the stressor and the post-stress sample resulted in larger effect sizes for IL-2, IL-4, TNF-α, but smaller effect sizes for IL-8. We urge further investigation of optimal post-stressor time points for salivary biomarkers, as studies utilizing blood samples show that sample timing may affect the magnitude of effects observed (12), and optimal sample timing may also vary by biomarker and length of stressor. For example, in post-hoc analyses we found that longer stressors yielded larger effect sizes for IL-2, IL-4 and TNF-α, suggesting that researchers may consider using longer stressors to elicit larger responses. Contrary to hypotheses, study quality was not a significant moderator of any biomarker assessed. The lack of findings linked to study quality was likely due a restriction of possible values, as most of our studies were deemed of moderate quality. Assay technique emerged as a significant moderator of effects for IL-8, with multiplex assays resulting in larger effect sizes for IL-8. However, of note, many reviewed studies still had significant limitations (e.g., there was a limited variability in mean age across many studies, and few studies assessed flow rate) and likely limited power. As such, we sought to not over-interpret our moderation results, and urge further study of potential moderators of stress-related changes in salivary markers of inflammation.

4.3 Limitations and Future Directions

The present systematic review and meta-analysis does have some limitations warranting more research, in addition to the suggestions listed above. First, results for unpublished studies were not obtained, which may contribute to “file-drawer error” (i.e., non-significant results are more likely to go unpublished, and as such, our results may be more biased towards statistical significance). In addition, many studies took multiple post-stress samples and the present study used the largest positive effect size. Although this approach helped reduce Type II error and was in keeping with our primary research question of whether salivary inflammatory markers increase with acute stress, it does introduce positive result bias. Future work should employ different approaches to answer questions that remain, including nuances around stress reactivity and recovery of salivary inflammatory markers across time. Additionally, as discussed above, some moderators could not be assessed due to missing data. Future research and meta-analyses would be greatly facilitated by open access of data so that researchers could explore moderators that may not have been part of the original focus of the published studies.

There were too few studies to enable a robust test of all our moderators; however, all moderation results for biomarkers with ≥ 5 samples are provided in supplemental tables to guide future research. Reviews in blood (47) and saliva (13) recommend considering exercise, smoking, medication use, menstrual cycle, age, gender, socioeconomic status, ethnicity, BMI, alcohol use, medication use, and sleep as potential confounds that may influence salivary and systemic inflammatory markers. These biobehavioral factors should

be assessed and reported whenever possible to guide future research. Research on this topic may also benefit from conducting studies designed to test specific moderators of interest (e.g., recruiting samples of age-matched men and women to test the effect of gender on stress reactivity).

Another related limitation was our inability to tease apart effects of anticipation from the stressor itself; future research should consider the importance of resting baseline samples. Such samples allow the participant to habituate to the novelty of the situation before stressor exposure and may allow for better determination of whether the stressor itself (vs. individual differences in anticipatory stress) is eliciting changes in biomarker levels.

The present review only included within-subjects designs (i.e., comparing participants to their own baseline). When studies did include a non-stressor control group, these results were excluded from our analyses. However, to ensure the validity of our results, we coded for whether studies used a separate manipulation check variable (i.e., subjective stress or an additional objective biomarker [heart rate, cortisol, blood pressure]); 88.1% of our studies used a manipulation check. Additionally, our quality guide included a question about whether the stressor used had been validated to induce stress in previous research, and just over half of our studies used a validated stressor. It would be valuable for future studies to incorporate cross-over designs, where the same individuals are exposed to both stressor and non-stressor conditions. This design would allow for inferences about whether observed effects were truly the result of stress exposure, opposed to individual differences.

Further investigation into the mediators and moderators of stress reactivity may elucidate pathways by which stress leads to changes in salivary inflammation and provide targets for intervention. For example, cognitive control of emotional material has been shown to moderate stress reactivity (48). Many studies in the current review also considered the role of state emotion, with negative emotion predicting salivary inflammation in some (43), but not all studies (37,43). Levels of positive emotion, perceived stress, and emotion regulation strategy use (e.g., rumination) also were associated with some salivary markers (40,43,46). Other physiological markers (e.g., heart rate variability, blood pressure, cortisol, or sympathetic activation) that are sometimes included in such research (often as manipulation checks) may offer additional insight into other mechanisms of salivary inflammation (49).

4.4 Potential Clinical Implications

Our intention for this meta-analysis was to provide a synthesis of findings related to salivary markers of inflammation and how they relate to acute stress, in hopes of expanding understanding in this field with this review. Blood-based markers are the current gold-standard for predicting peripheral disease processes and the present review is unable to evaluate the feasibility of using salivary markers as a replacement. Some studies have shown that cytokines such as IL-6, IL-10, IL-18, INF-γ, and IL-1β demonstrate modest correlations between blood and saliva (r = .29–59) (19,20,50,51). However, other studies have reported null associations between salivary and blood-based inflammatory markers (19,52). These mixed findings suggest more work is needed to determine the extent to which salivary inflammatory markers correspond to blood-based inflammation markers.

Regardless of how salivary markers map onto blood, they may provide important information about disease processes. Salivary markers appear to have some predictive utility for diseases such as cardiovascular disease, rheumatoid arthritis, and metabolic syndrome (22–25,53). Some inflammatory biomarkers also show cross-sectional associations with mental health conditions such as PTSD (20,54), with preliminary evidence suggesting that childhood (49) and lifetime trauma (40) may be associated with increases in salivary inflammation in response to acute stress. Together, these studies suggest that salivary inflammatory markers may provide an important local measure of inflammation, and one which may serve as an important biomarker of health. The degree of change and level of recovery in inflammatory markers in response to acute stress also is likely indicative of individual differences in how people respond to stressors in everyday life and may reflect vulnerability to disease (12). Further work assessing both salivary and blood-based markers of inflammation in response to stress is needed, both to understand their correspondence and utility for understanding disease risk.

Our findings demonstrate that multiple salivary inflammatory biomarkers reliably increase in response to stress. As a next step, researchers could examine if this reactivity predicts distal health outcomes. Should predictive validity be established, salivary measures of inflammation may hold potential to inform questions related to the effects of stress in disease. The non-invasive nature and low cost of salivary measures may also enable studies using multiple daily assessments in naturalistic settings. Such research is needed to understand how daily stress can translate to acute physiological changes and vulnerability to stress-related disease across time.

5. Conclusions

Our pre-registered systematic review and meta-analysis provides the first quantitative summary of changes in salivary inflammatory markers in response to acute stress. Cytokines IL-6, IL-10, TNF-α, and INF-γ significantly increased in response to stress, exhibiting moderate to large effect sizes. Post hoc sensitivity analyses revealed that IL-1β and IL-8 also increased post-stress but findings with INF-γ did not hold after removing outliers. Examination of moderators suggested that study methodology and sample demographics moderated some associations; in general, however, moderation results were inconsistent across studies and biomarkers. This synthesis of the burgeoning literature on salivary markers of inflammation suggests that salivary biomarkers may serve as useful markers of stress reactivity in biobehavioral research. Additional research is warranted to tease apart potential mediators and moderators of these associations, as well as understand the relation of inflammatory measured in saliva to other mediums (e.g., blood) and the predictive utility of salivary inflammation for health outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

[Author's] work was supported in part by the Department of Veterans Affairs Office of Academic Affiliations Advanced Fellowship Program in Mental Illness Research and Treatment, the Research Service of the Central Texas Veterans Health Care System, and the VISN 17 Center of Excellence for Research on Returning War Veterans. [Author's] work was also supported in part by Career Development Award Number IK1 RX003122 from the United States (U.S.) Department of Veterans Affairs Rehabilitation R&D (Rehab RD) Service. [Author's work] is supported by grants NIH/NIAID R01AI128359-01 and DoD-VA 1I01CU000144-0. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs, United States government, or academic affiliates, nor were these sources involved in study design, data collection, or interpretation of findings. These funding sources were not involved in study design, data collection, or interpretation of findings. The authors would like to thank Hailey Fox and Abigail Herschbach for their assistance with verifying reporting of results and manuscript formatting.

References

1. Kiecolt-Glaser JK, Derry HM, Fagundes CP. Inflammation: Depression fans the flames and feasts on the heat. Am J Psychiatry. 2015; DOI: 10.1176/appi.ajp.2015.15020152
2. Del Giudice M, Gangestad SW. Rethinking IL-6 and CRP: Why they are more than inflammatory biomarkers, and why it matters. Brain Behav Immun. 2018; DOI: 10.1016/j.bbi.2018.02.013
3. Cunningham C, Campion S, Lunnon K, Murray CL, Woods JFC, Deacon RMJ, et al. Systemic Inflammation Induces Acute Behavioral and Cognitive Changes and Accelerates Neurodegenerative Disease. Biol Psychiatry. 2009; DOI: 10.1016/j.biopsych.2008.07.024
4. Dowlati Y, Herrmann N, Swardfager W, Liu H, Sham L, Reim EK, et al. A Meta-Analysis of Cytokines in Major Depression. Biol Psychiatry. 2010; DOI: 10.1016/j.biopsych.2009.09.033
5. Il'yasova D, Colbert LH, Harris TB, Newman AB, Bauer DC, Satterfield S, et al. Circulating levels of inflammatory markers and cancer risk in the health aging and body composition cohort. Cancer Epidemiol Biomarkers Prev. 2005; DOI: 10.1158/1055-9965.EPI-05-0316
6. Kaspersen KA, Dinh KM, Erikstrup LT, Burgdorf KS, Pedersen OB, Sørensen E, et al. Low-grade inflammation is associated with susceptibility to infection in healthy men: Results from the Danish Blood Donor study (DBDS). PLoS One. 2016; DOI: 10.1371/journal.pone.0164220
7. Kiecolt-Glaser JK, Loving TJ, Stowell JR, Malarkey WB, Lemeshow S, Dickinson SL, et al. Hostile marital interactions, proinflammatory cytokine production, and wound healing. Arch Gen Psychiatry. 2005; DOI: 10.1001/archpsyc.62.12.1377
8. Michopoulos V, Powers A, Gillespie CF, Ressler KJ, Jovanovic T. Inflammation in Fear-and Anxiety-Based Disorders: PTSD, GAD, and beyond. Neuropsychopharmacology. 2017 DOI: 10.1038/npp.2016.146
9. Passos IC, Vasconcelos-Moreno MP, Costa LG, Kunz M, Brietzke E, Quevedo J, et al. Inflammatory markers in post-traumatic stress disorder: A systematic review, meta-analysis, and meta-regression. The Lancet Psychiatry. 2015; DOI: 10.1016/S2215-0366(15)00309-0
10. Ridker PM, Hennekens CH, Buring JE, Rifai N. C-Reactive Protein and Other Markers of Inflammation in the Prediction of Cardiovascular Disease in Women. N Engl J Med. 2000; DOI: 10.1056/NEJM200003233421202
11. Steptoe A, Hamer M, Chida Y. The effects of acute psychological stress on circulating inflammatory factors in humans: A review and meta-analysis. Brain, Behavior, and Immunity. 2007 DOI: 10.1016/j.bbi.2007.03.011
12. Marsland AL, Walsh C, Lockwood K, John-Henderson NA. The effects of acute psychological stress on circulating and stimulated inflammatory markers: A systematic review and meta-analysis. Brain Behav Immun. 2017; DOI: 10.1016/j.bbi.2017.01.011
13. Slavish DC, Graham-Engeland JE, Smyth JM, Engeland CG. Salivary markers of inflammation in response to acute stress. Brain, Behavior, and Immunity. 2015 DOI: 10.1016/j.bbi.2014.08.008
14. Cohen S, Hamrick N. Stable individual differences in physiological response to stressors: Implications for stress-elicited changes in immune related health. Brain, Behavior, and Immunity. 2003 DOI: 10.1016/S0889-1591(03)00110-7
15. Marsland AL, Bachen EA, Cohen S, Rabin B, Manuck SB. Stress, immune reactivity and susceptibility to infectious disease. Physiol Behav. 2002; DOI: 10.1016/S0031-9384(02)00923-X

16. Smyth J, Zawadzki M, Gerin W. Stress and disease: A structural and functional analysis. Soc Personal Psychol Compass. 2013; DOI: 10.1111/spc3.12020
17. Desai GS, Mathews ST. Saliva as a non-invasive diagnostic tool for inflammation and insulin-resistance. World J Diabetes. 2014; DOI: 10.4239/wjd.v5.i6.730
18. Engeland CG, Bosch JA, Rohleder N. Salivary biomarkers in psychoneuroimmunology. Curr Opin Behav Sci. 2019;28:58–65. [PubMed: 32215283]
19. La Fratta I, Tatangelo R, Campagna G, Rizzuto A, Franceschelli S, Ferrone A, et al. The plasmatic and salivary levels of IL-1β, IL-18 and IL-6 are associated to emotional difference during stress in young male. Sci Rep. 2018; DOI: 10.1038/s41598-018-21474-y
20. Newton TL, Fernandez-Botran R, Miller JJ, Burns VE. Interleukin-6 and soluble interleukin-6 receptor levels in posttraumatic stress disorder: Associations with lifetime diagnostic status and psychological context. Biol Psychol. 2014; DOI: 10.1016/j.biopsycho.2014.03.009
21. Williamson S, Munro C, Pickler R, Grap MJ, Elswick RK. Comparison of Biomarkers in Blood and Saliva in Healthy Adults. Nurs Res Pract. 2012; DOI: 10.1155/2012/246178
22. Ebersole JL, Kryscio RJ, Campbell C, Kinane DF, McDevitt J, Christodoulides N, et al. Salivary and serum adiponectin and C-reactive protein levels in acute myocardial infarction related to body mass index and oral health. J Periodontal Res. 2017; DOI: 10.1111/jre.12406
23. Gohel V, Jones JA, Wehler CJ. Salivary biomarkers and cardiovascular disease: A systematic review. Clin Chem Lab Med. 2018; DOI: 10.1515/cclm-2017-1018
24. Chauhan A, Yadav SS, Dwivedi P, Lal N, Usman K, Khattri S. Correlation of Serum and Salivary Cytokines Level With Clinical Parameters in Metabolic Syndrome With Periodontitis. J Clin Lab Anal. 2016; DOI: 10.1002/jcla.21917
25. Silvestre-Rangil J, Bagán L, Silvestre FJ, Martinez-Herrera M, Bagán J. Periodontal, salivary and IL-6 status in rheumatoid arthritis patients: A cross-sectional study. Med Oral Patol Oral Cir Bucal. 2017;22(5):e595–600. DOI: Doi:10.4317/medoral.21937 [PubMed: 28809379]
26. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015; DOI: 10.1186/2046-4053-4-1
27. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (prisma-p) 2015: Elaboration and explanation. BMJ (Online). 2015 DOI: 10.1136/bmj.g7647
28. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: A proposal for reporting. J Am Med Assoc. 2000; DOI: 10.1001/jama.283.15.2008
29. Slavish DC, Szabo YZ. The effect of acute stress on salivary markers of inflammation: A systematic review protocol. Syst Rev. 2019;108(8):1–8. DOI: 10.1186/s13643-019-1026-4
30. Brandtzaeg P Do salivary antibodies reliably reflect both mucosal and systemic immunity? In: Annals of the New York Academy of Sciences. 2007 DOI: 10.1196/annals.1384.012
31. 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; DOI: 10.1111/j.0006-341X.2000.00455.x
32. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front Psychol. 2013;6:863.
33. Hirst RJ, Cragg L, Allen HA. Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. Neuroscience and Biobehavioral Reviews. 2018 DOI: 10.1016/j.neubiorev.2018.07.012
34. Borenstein M, Hedges L, Higgins J, Rothstein H. Introduction to meta-analysis. John Wiley & Sons; 2011.
35. Higgins J, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. 2011 section 9.6.5.1.
36. Viechtbauer W Conducting Meta-Analyses in R with the metafor Package. J Stat Softw. 2010; DOI: 10.18637/jss.v036.i03
37. Moreira A, Bacurau RFP, Napimoga MH, Arruda AFS, Freitas CG, Drago G, et al. Salivary IL-21 and IGA responses to a competitive match in elite basketball players. Biol Sport. 2013; DOI: 10.5604/20831862.1077548

38. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008; DOI: 10.1136/bmj.39489.470347.ad
39. Slavish DC, Jones DR, Smyth JM, Engeland CG, Song S, McCormick NM, et al. Positive and Negative Affect and Salivary Markers of Inflammation Among Young Adults. Int J Behav Med. 2019;1–12. [PubMed: 30725353]
40. Szabo YZ, Fernandez-Botran R, Newton TL. Cumulative trauma, emotion reactivity and salivary cytokine levels following acute stress in healthy women. Anxiety, Stress Coping. 2019; DOI: 10.1080/10615806.2018.1524377
41. Szabo YZ, Newton TL, Miller JJ, Lyle KB, Fernandez-Botran R. Acute stress induces increases in salivary IL-10 levels. Stress. 2016; DOI: 10.1080/10253890.2016.1206885
42. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. BMJ. 2009; DOI: 10.1136/bmj.b2393
43. Newton TL, Fernandez-Botran R, Lyle KB, Szabo YZ, Miller JJ, Warnecke AJ. Salivary cytokine response in the aftermath of stress: An emotion regulation perspective. Emotion. 2017; DOI: 10.1037/emo0000156
44. Weidner G, Helmig L. Cardiovascular stress reactivity and mood during the menstrual cycle. Women Heal. 1990; DOI: 10.1300/J013v16n03_02
45. Shields GS, Dunn TM, Trainor BC, Yonelinas AP. Determining the biological associates of acute cold pressor post-encoding stress effects on human memory: The role of salivary interleukin-1β. Brain Behav Immun. 2019; DOI: 10.1016/j.bbi.2019.06.011
46. Buzgoova K, Balagova L, Marko M, Kapsdorfer D, Riecansky I, Jezova D. Higher perceived stress is associated with lower cortisol concentrations but higher salivary interleukin-1beta in socially evaluated cold pressor test. Stress. 2019; DOI: 10.1080/10253890.2019.1660872
47. O'Connor MF, Bower JE, Cho HJ, Creswell JD, Dimitrov S, Hamby ME, et al. To assess, to control, to exclude: Effects of biobehavioral factors on circulating inflammatory markers. Brain Behav Immun. 2009; DOI: 10.1016/j.bbi.2009.04.005
48. Shields GS, Kuchenbecker SY, Pressman SD, Sumida KD, Slavich GM. Better cognitive control of emotional information is associated with reduced pro-inflammatory cytokine reactivity to emotional stress. Stress. 2016; DOI: 10.3109/10253890.2015.1121983
49. Tell D, Mathews HL, Burr RL, Janusek LW. During stress, heart rate variability moderates the impact of childhood adversity in women with breast cancer. Stress. 2018; DOI: 10.1080/10253890.2018.1424132
50. Dan H, Liu W, Wang J, Wang Z, Wu R, Chen Q, et al. Elevated IL-10 concentrations in serum and saliva from patients with oral lichen planus. Quintessence Int (Berl). 2011;
51. Fernandez-Botran R, Miller JJ, Burns VE, Newton TL. Correlations among inflammatory markers in plasma, saliva and oral mucosal transudate in post-menopausal women with past intimate partner violence. Brain Behav Immun. 2011; DOI: 10.1016/j.bbi.2010.09.023
52. Cullen T, Thomas AW, Webb R, Hughes MG. The relationship between interleukin-6 in saliva, venous and capillary plasma, at rest and in response to exercise. Cytokine. 2015; DOI: 10.1016/j.cyto.2014.10.011
53. Out D, Hall RJ, Granger DA, Page GG, Woods SJ. Assessing salivary C-reactive protein: Longitudinal associations with systemic inflammation and cardiovascular disease risk in women exposed to intimate partner violence. Brain Behav Immun. 2012; DOI: 10.1016/j.bbi.2012.01.019
54. Wang Z, Mandel H, Levingston CA, Young MRI. An exploratory approach demonstrating immune skewing and a loss of coordination among cytokines in plasma and saliva of Veterans with combat-related PTSD. Hum Immunol. 2016; DOI: 10.1016/j.humimm.2016.05.018
55. Acurio J, Celis C, Perez J, Escudero C. Acoustic parameters and salivary IL-6 levels in overweight and obese teachers. J Voice. 2014; DOI: 10.1016/j.jvoice.2014.03.002
56. Auer BJ, Calvi JL, Jordan NM, Schrader D, Byrd-Craven J. Communication and social interaction anxiety enhance interleukin-1 beta and cortisol reactivity during high-stakes public speaking. Psychoneuroendocrinology. 2018; DOI: 10.1016/j.psyneuen.2018.05.011

57. Brandenburg JP, Giles LV. Four Days of Blueberry Powder Supplementation Lowers the Blood Lactate Response to Running But Has No Effect on Time-Trial Performance. Int J Sport Nutr Exerc Metab. 2019; DOI: 10.1123/ijsnem.2019-0040
58. Caris AV, da Silva ET, dos Santos SA, Tufik S, dos Santos RVT. Effects of carbohydrate and glutamine supplementation on oral mucosa immunity after strenuous exercise at high altitude: A double-blind randomized trial. Nutrients. 2017; DOI: 10.3390/nu9070692
59. Chiappelli J, Shi Q, Kodi P, Savransky A, Kochunov P, Rowland LM, et al. Disrupted glucocorticoid-Immune interactions during stress response in schizophrenia. Psychoneuroendocrinology. 2016; DOI: 10.1016/j.psyneuen.2015.09.010
60. Daglius Dias R, Scalabrini Neto A. Stress levels during emergency care: A comparison between reality and simulated scenarios. J Crit Care. 2016; DOI: 10.1016/j.jcrc.2016.02.010
61. Dogra S, Wolf M, Jeffrey MP, Foley RCA, Logan-Sprenger H, Jones-Taggart H, et al. Disrupting prolonged sitting reduces IL-8 and lower leg swell in active young adults. BMC Sports Sci Med Rehabil. 2019; DOI: 10.1186/s13102-019-0138-4
62. Ersche KD, Hagan CC, Smith DG, Abbott S, Jones PS, Apergis-Schoute AM, et al. Aberrant disgust responses and immune reactivity in cocaine-dependent men. Biol Psychiatry. 2014; DOI: 10.1016/j.biopsych.2013.08.004
63. Filaire E, Larue J, Portier H, Abed A, Graziella PD, Teixeira A, et al. Lecturing to 200 students and its effects on cytokine concentration and salivary markers of adrenal activation. Stress Heal. 2011; DOI: 10.1002/smi.1332
64. Gassen J, Makhanova A, Maner JK, Plant EA, Eckel LA, Nikonova L, et al. Experimentally-Induced Inflammation Predicts Present Focus. Adapt Hum Behav Physiol. 2019; DOI: 10.1007/s40750-019-00110-7
65. Groer M, Murphy R, Bunnell W, Salomon K, Van Eepoel J, Rankin B, et al. Salivary measures of stress and immunity in police officers engaged in simulated critical incident scenarios. J Occup Environ Med. 2010; DOI: 10.1097/JOM.0b013e3181e129da
66. Izawa S, Sugaya N, Kimura K, Ogawa N, Yamada KC, Shirotsuki K, et al. An increase in salivary interleukin-6 level following acute psychosocial stress and its biological correlates in healthy young adults. Biol Psychol. 2013; DOI: 10.1016/j.biopsycho.2013.06.006
67. Kennedy PJ, Cryan JF, Quigley EMM, Dinan TG, Clarke G. A sustained hypothalamic-pituitary-Adrenal axis response to acute psychosocial stress in irritable bowel syndrome. Psychol Med. 2014; DOI: 10.1017/S003329171400052X
68. Kimura K, Izawa S, Sugaya N, Ogawa N, Yamada KC, Shirotsuki K, et al. The biological effects of acute psychosocial stress on delay discounting. Psychoneuroendocrinology. 2013; DOI: 10.1016/j.psyneuen.2013.04.019
69. Laurent HK, Lucas T, Pierce J, Goetz S, Granger DA. Coordination of cortisol response to social evaluative threat with autonomic and inflammatory responses is moderated by stress appraisals and affect. Biol Psychol. 2016; DOI: 10.1016/j.biopsycho.2016.04.066
70. Maydych V, Claus M, Watz C, Kleinsorge T. Attention to emotional information is associated with cytokine responses to psychological stress. Front Neurosci. 2018; DOI: 10.3389/fnins.2018.00687
71. Melo BP, Guariglia DA, Pedro RE, Bertolini DA, de Paula Ramos S, Peres SB, et al. Combined Exercise Modulates Cortisol, Testosterone, and Immunoglobulin A Levels in Individuals Living With HIV/AIDS. J Phys Act Heal. 2019; DOI: 10.1123/jpah.2019-0134
72. Minetto M, Rainoldi A, Gazzoni M, Terzolo M, Borrione P, Termine A, et al. Differential responses of serum and salivary interleukin-6 to acute strenuous exercise. Eur J Appl Physiol. 2005; DOI: 10.1007/s00421-004-1241-z
73. Minetto MA, Gazzoni M, Lanfranco F, Baldi M, Saba L, Pedrola R, et al. Influence of the sample collection method on salivary interleukin-6 levels in resting and post-exercise conditions. Eur J Appl Physiol. 2007; DOI: 10.1007/s00421-007-0484-x
74. Quinn ME, Stanton CH, Slavich GM, Joormann J. Executive Control, Cytokine Reactivity to Social Stress, and Depressive Symptoms: Testing the Social Signal Transduction Theory of Depression. Stress. 2019; DOI: 10.1080/10253890.2019.1641079
75. Reed RG, O'Connor MF, Pace TWW, Raison CL, Butler EA. Dyadic coping and salivary interleukin-6 responses to interpersonal stress. J Fam Psychol. 2017; DOI: 10.1037/fam0000249

76. Rodríguez-Medina DA, Leija-Alva G, Domínguez-Trejo B, Hernández-Pozo M del R, Cruz-Albarrán IA, Morales-Hernández LA, et al. Effects of the Trier Social Stress Test on the distributions of IL-6 and MAP levels. Heliyon. 2019; DOI: 10.1016/j.heliyon.2019.e01580

77. Saban KL, Mathews HL, Bryant FB, Tell D, Joyce C, DeVon HA, et al. Perceived discrimination is associated with the inflammatory response to acute laboratory stress in women at risk for cardiovascular disease. Brain Behav Immun. 2018; DOI: 10.1016/j.bbi.2018.07.010

78. Usui T, Yoshikawa T, Ueda SY, Katsura Y, Orita K, Fujimoto S. Effects of acute prolonged strenuous exercise on the salivary stress markers and inflammatory cytokines. Japanese J Phys Fit Sport Med. 2011; DOI: 10.7600/jspfsm.60.295

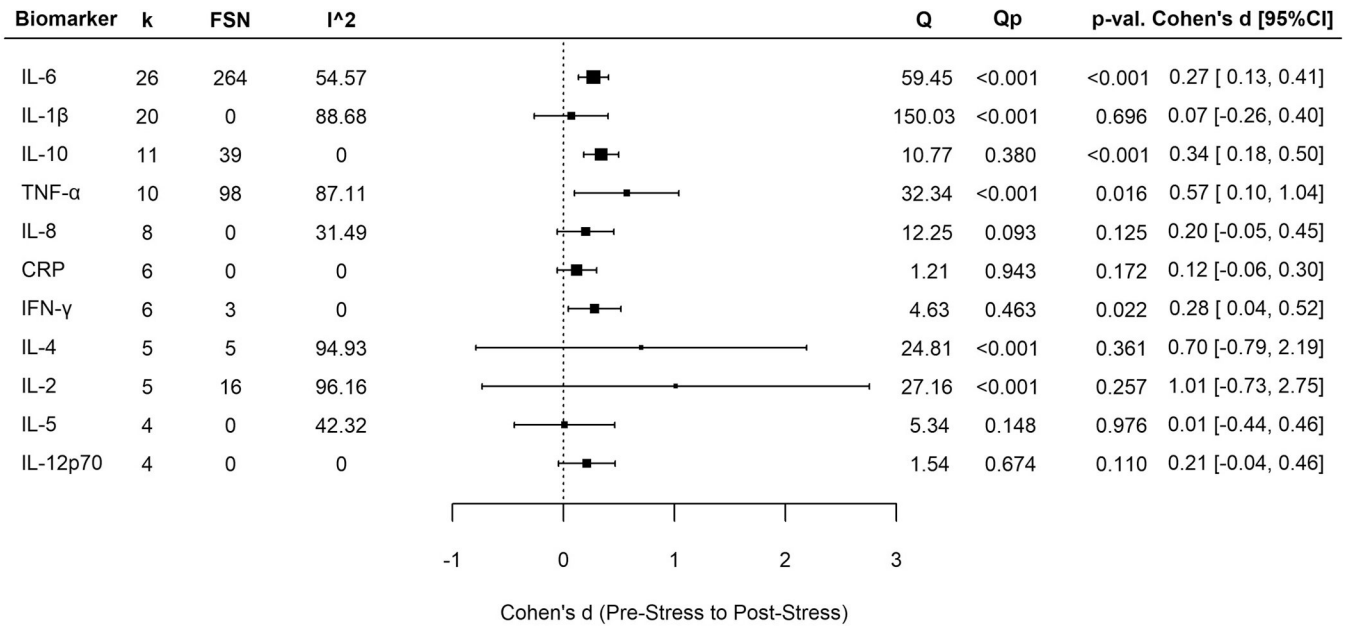

Figure 1.

Forest plot displaying omnibus effect size for each biomarker included in meta-analyses. IL = interleukin, TNF = tumor necrosis factor, CRP = C-reactive protein, IFN = interferon, FSN = Fail-safe N (number of missing studies averaging a z-value of zero that should be added to make the combined effect size statistically insignificant), k = number of studies included in each meta-analysis, Î2 = percentage of variation across studies that is due to heterogeneity rather than chance, Q = Cochran's Q (heterogeneity statistic, weighted sum of squared differences between individual study effects and the pooled effect across studies), Qp = p-value of Q (p-values < .05 indicate significant heterogeneity), p-value = overall p-value of omnibus meta-analysis, Cohen's dav = Cohen's d with standard deviations averaged across the two time points (see equation 10 in Lakens 2013) (23), 95% CI = 95% confidence interval of omnibus Cohen's dav.

Table 1

Study Characteristics

Citation	original N	analytic N	% women	% White	Sample Health Status	Biomarkers assessed	Saliva Collection Method	Stressor	Stressor Length (in minutes)	Time since start of stressor samples taken (in minutes)	Results summary	Total Quality Rating
Acurio et al., 2014 (sample 2) (55)	8	8	50%	not reported	Healthy *	IL-6	passive drool	speech	60–90	0, unclear	IL-6: **d_1 = −1.25**	3
Auer et al., 2018 (56)	95	66–67	61.05%	66.70%	Healthy	IL-1β	passive drool	speech	7	0, 7, 27	IL-1β: **d_1 = −.09**, d_2 = −.19	9
Brandenburg et al., 2019 (57) †	14	10–13	38.5%	not reported	Healthy	CRP	passive drool	exercise	50.2	0, 56.2	CRP: **d_1 =.20**	9
Buzgoova et al., 2019 (46)	24	23–24	0%	100%	Healthy	IL-1β	Salivette/ cotton	other (socially evaluated cold pressor test)	2	0, 2, 7, 17, 32	IL-1β: d_1 = −.67, d_2 = −.53, d_3 = **−.43**, d_4 = −.49	9
Caris et al., 2017 (58) ‡	15	15	0%	not reported	Healthy	IL-6, IL-10 TNF-α	Salivette/ cotton	exercise	27.6	0, 27.6, 147.6	IL-6: d_1 = −.17, **d_2** = **−.28** IL-10: **d_1 = −.28**, d_2 = −.05 TNF-α: d_1 = −.09, **d_2 = −.38**	7
Chiappelli et al., 2016 (sample 1) (59)	34	34	41.18%	44.10%	Clinical (Schizophrenia)	IL-6	passive drool	cognitive	14	0, 14, 34, 54	IL-6: **d_1 = .30**, d_2 = −.19, d_3 = −.05	10
Chiappelli et al., 2016 (sample 2) (59)	40	40	50%	47.50%	Healthy	IL-6	passive drool	cognitive	14	0, 14, 34, 54	IL-6: d_1 = .31, **d_2** = **.34**, d_3 = .20	11
Dias et al., 2016 (60)	16	16	56.25%	not reported	Healthy	IL-1β	Salivette/ cotton	other (simulated emergency situations)	25	0, 25	IL-1β: **d_1 = .93**	7
Dogra et al., 2019 (61)	10	10	50%	80%	Healthy	IL-1β, IL-8,	Salivette/ cotton	exercise	9	0, 240	IL-1β: **d_1 = −1.09** IL-8: **d_1 = −1.12**	12
Ersche et al., 2014 (sample 1) (62)	31	31	0%	not reported	Clinical (Cocaine Use Disorder)	IFN-γ, IL-1β, IL-6, IL-8, IL-10,	passive drool	other (disgusting images)	not reported	unclear	IFN-γ: **d_1 = .63** IL-1β: **d_1 = .75** IL-6: **d_1 = .34**	7

Citation	original N	analytic N	% women	% White	Sample Health Status	Biomarkers assessed	Saliva Collection Method	Stressor	Stressor Length (in minutes)	Time since start of stressor samples taken (in minutes)	Results summary	Total Quality Rating
						IL-12p70, TNF-α					IL-8: **d_1 = .73** IL-10: **d_1 = .57** IL-12p70: **d_1 = .42** TNF-α: **d_1 = .61**	
Ersche et al., 2014 (sample 2) (62)	30	30	0%	not reported	Mixed	IFN-γ, IL-1β, IL-6, IL-8, IL-10, IL-12p70, TNF-α	passive drool	other (disgusting images)	not reported	unclear	IFN-γ: **d_1 = .35** IL-1β: **d_1 = .45** IL-6: **d_1 = .21** IL-8: **d_1 = .19** IL-10: **d_1 = .45** IL-12p70: **d_1 = .30** TNF-α: **d_1 = .20**	7
Filaire et al., 2009 (63)	9	9	22.22%	not reported	Healthy	IL-2, IL-4, IL-10, TNF-α	Salivette/ cotton	speech	120	0, 120, 240	IL-2: d_1 = .82, **d_2 = 5.08** IL-4: d_1 = .79, **d_2 = 4.22** IL-10: **d_1 = .22**, d_2 = −.33 TNF-α: d_1 = −.26, **d_2 = 2.96**	7
Gassen et al., 2019 (sample 1) (64)	54	IL-1β: 45–47 IL-6: 19–23	66.67%	74%	Healthy	IL-1β, IL-6	passive drool	other (“pathogen threat”; images of physical illness)	5	0, 25	IL-1β: **d_1 = .27** IL-6: **d_1 = .34**	7
Gassen et al., 2019 (sample 2) (64)	53	IL-1β: 44–45 IL-6: 31–32	86.79%	75%	Healthy	IL-1β, IL-6	passive drool	other (“physical harm threat”; images of violence or other forms of physical harm	5	0, 25	IL-1β: **d_1 = .13** IL-6: **d_1 = .45**	7
Groer et al., 2010 (sample 1) (65)	49	48–49	22.45%	70.40%	not reported	IL-6	passive drool	other (virtual reality critical incident, chase)	2	0, 12, 32	IL-6: d_1 = −.22, **d_2 = .02**	7
Groer et al., 2010 (sample 2) (65)	92	96–98	6.52%	83%	not reported	IL-6	passive drool	other (virtual reality critical incident, workplace shooter)	6	0, 16, 36	IL-6: d_1 = .16, **d_2 = .18**	7
Izawa et al., 2013 (66)	50	50	22%	not reported	Healthy	IL-6	passive drool	TSST	20	0, 10, 15, 20, 30, 40, 50, 65, 80	IL-6: **d_1 = .47**, d_2 = .42, d_3 = .38, d_4 = .35, d_5 = .29, d_6	10

Citation	original N	analytic N	% women	% White	Sample Health Status	Biomarkers assessed	Saliva Collection Method	Stressor	Stressor Length (in minutes)	Time since start of stressor samples taken (in minutes)	Results summary	Total Quality Rating
											= .19, d_7 = .29, d_8 = .25	
Kennedy et al., 2014 (sample 1) (67)	13	13	100%	not reported	Clinical (Irritable Bowel Syndrome)	CRP	Salivette/ cotton	TSST	20	0, 20, 35, 65	CRP: $\mathbf{d_1}$ = **.14**, d_2 = −.46, d_3 = −.05	11
Kennedy et al., 2014 (sample 2) (67)	15	15	100%	not reported	Healthy	CRP	Salivette/ cotton	TSST	20	0, 20, 35, 65	CRP: $\mathbf{d_1}$ = **.02**, d_2 = −.43, d_3 = −.40	11
Kimura et al., 2013 (68)	39	26	not reported	not reported	Healthy	CRP	spitting	TSST	20	0, 10, 15, 20, 30, 40, 50, 65, 80	CRP: d_1 = .18, $\mathbf{d_2}$ = **.23**, d_3 = .14, d_4 = .01, d_5 = .04, d_6 = −.04, d_7 = .02, d_8 = −.05	8
La Fratta et al., 2018 (19)	61	61	0%	100%	Healthy	CRP, IL-1β, IL-6, IL-18	Salivette/ cotton	other (academic examination)	30	−30, 60	CRP: $\mathbf{d_1}$ = **−.03** IL-1β: $\mathbf{d_1}$ = **−1.84** IL-6: $\mathbf{d_1}$ = **1.27** IL-18: $\mathbf{d_1}$ = **−0.89**	10
Laurent et al., 2016 (69)	115	115	69.57%	0%	Mixed	CRP	passive drool	TSST	20	unclear	CRP: d_1 = .17, d_2 = .18, d_3 = .09, d_4 = .13, $\mathbf{d_5}$ = **.18**	12
Maydych et al., 2018 (sample 1) (70)	30	30	50%	not reported	Healthy	IFN-α, IFN-γ, IL-1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12p70, IL-13, IL-17A, TNF-α	Salivette/ cotton	other (PASAT + social evaluative stress)	25	−5, 30, 70	IFN-α: d_1 = −.04, $\mathbf{d_2}$ = **.33** IFN-γ: d_1 = −.04, $\mathbf{d_2}$ = **.31** IL-1β: d_1 =−.05, $\mathbf{d_2}$ = **.24** IL-2: d_1 =.01, $\mathbf{d_2}$ = **.33** IL-4: d_1 = .12, $\mathbf{d_2}$ = **.26** IL-5: d_1 = .05, $\mathbf{d_2}$ = **.36** IL-6: d_1 =.003, $\mathbf{d_2}$ = **.12** IL-8: d_1 =−.02, $\mathbf{d_2}$ = **.18** IL-10: d_1 = −.06, $\mathbf{d_2}$ = **.28** IL-12p70: d_1 = −.04, $\mathbf{d_2}$ = **.04**	11

Citation	original N	analytic N	% women	% White	Sample Health Status	Biomarkers assessed	Saliva Collection Method	Stressor	Stressor Length (in minutes)	Time since start of stressor samples taken (in minutes)	Results summary	Total Quality Rating
											IL-13: d_1 = .12, **d_2 = .39** IL-17A: d_1 = −.07, **d2 = −.07** TNF-α: d_1 = −.10, **d_2 = .40**	
Maydych et al., 2018 (sample 2) (70)	30	30	50%	not reported	Healthy	IFN-α, IFN-γ, IL-1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, IL-12p70, IL-13, IL-17A, TNF-α	Salivette/ cotton	cognitive	25	−5, 30, 70	IFN-α: d_1 = −.16, **d_2 = .08** IFN-γ: d_1 = −.16, **d_2 = .05** IL-1β: d_1 =−.05, **d_2 = .001** IL-2: d_1 =−.11, **d_2 = .14** IL-4: d_1 = −.18, **d_2 = .03** IL-5: d_1 = −.08, **d_2 = .16** IL-6: d_1 =−.15, **d_2 = −.14** IL-8: d_1 = .24, **d_2 = .40** IL-10: d_1 = −.18, **d_2 = .13** IL-12p70: d_1 = −.13, **d_2 = .06** IL-13: d_1 = −.11, **d_2 = .11** IL-17A: d_1 = −.10, **d_2 = −.01** TNF-α: d_1 = −.29, **d_2 = .15**	11
Melo et al., 2019 (sample 1) (71)	9	9	not reported	not reported	Clinical (HIV+)	GM-CSF, IFN-γ, IL-1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, TNF-α	Salivette/ cotton	exercise	60	0, 60	GM-CSF : **d_1 = .87** IFN-γ: **d_1 = .19** IL-1β: **d_1 = −.57** IL-2: **d_1 = .58** IL-4: **d_1 = −.07** IL-5: **d_1 = −.89** IL-6: **d_1 = −.34** IL-8: **d_1 = .05** IL-10: **d_1 = −.89**	6

Citation	original N	analytic N	% women	% White	Sample Health Status	Biomarkers assessed	Saliva Collection Method	Stressor	Stressor Length (in minutes)	Time since start of stressor samples taken (in minutes)	Results summary	Total Quality Rating
Melo et al., 2019 (sample 2) (71)	9	9	not reported	not reported	Mixed	GM-CSF, IFN-γ, IL-1β, IL-2, IL-4, IL-5, IL-6, IL-8, IL-10, TNF-α	Salivette/ cotton	exercise	60	0, 60	GM-CSF: **d_1 = −.30** IFN-γ: **d_1 = −.39** IL-1β: **d_1 = −.42** IL-2: **d_1 = −.40** IL-4: **d_1 = −.38** IL-5: **d_1 = −.17** IL-6: **d_1 = −.07** IL-8: **d_1 = .07** IL-10: **d_1 = .35**	6
Minetto et al., 2005 (sample 1) (72)	7	7	42.86%	not reported	not reported	IL-6	Salivette/ Cotton	exercise	180	−15, −5, 180	IL-6: **d_1 = 1.11**	6
Minetto et al., 2007 (73)[†]	15	15	0%	100%	Healthy	IL-6	spitting; Salivette/ cotton	exercise	40	0, 40, 70, 100, 130, 160	IL-6: **d_1 = .62**, d_2 = .48, d_3 = .35, d_4 = .41, d_5 = .52	5
Moreira et al., 2013 (37)	20	20	0%	not reported	not reported	IL-21	passive drool	exercise	150	0, 160	IL-21: d not reported	5
Newton et al., 2017 (sample 1) (43)	68	68	0%	63%	Mixed	IL-1β, IL-6, TNF-α	Salivette/ cotton	TSST	10	−10, 50	IL-1β: **d_1 = .16** IL-6: **d_1 = .14** TNF-α: **d_1 = .17**	12
Newton et al., 2017 (sample 2) (43)	68	68	72.06%	68%	Mixed	IL-1β, IL-6, TNF-α	Salivette/ cotton	other (angry autobiographical memory)	10	−10, 50	IL-1β: **d_1 = .44** IL-6: **d_1 = .37** TNF-α: **d_1 = .31**	11
Quinn et al., 2019 (74)	23	20	75%	50%	Mixed	IL-1β, IL-6, TNF-α	passive drool	TSST	15	−33.51, 44.74	IL-1β: **d_1 = −.04** IL-6: **d_1 = −.01** TNF-α: **d_1 = .29**	9
Reed et al., 2017 (75)	48	46–47	0%	73%	Healthy	IL-6	other (SalivaBio non cotton swab)	other (heated discussion with partner)	15	0, 15, 60, 135	IL-6: d_1 = .16, **d_2 = .26**, d_3 = −.39	11
Rodriguez-Medina et al., 2019 (76)	42	41	35.70%	not reported	Healthy	IL-6	Salivette/ cotton	TSST	13	0, 13	IL-6: **d_1 = 0.24**	9
Saban et al., 2018 (77)	99	98–99	70.59%	49.50%	Mixed	IL-6	Salivette/ cotton	TSST	10	0, 25, 40, 55, 70, 85, 100	IL-6: **d_1 = .17**, d_2 = .14, d_3 = .08, d_4	12

Citation	original N	analytic N	% women	% White	Sample Health Status	Biomarkers assessed	Saliva Collection Method	Stressor	Stressor Length (in minutes)	Time since start of stressor samples taken (in minutes)	Results summary	Total Quality Rating
											= −.01, d_5 = −.08, d_6 = −.13	
Shields et al., 2019 (45)	38	38	57.90%	21%	Healthy	IL-1β,	passive drool	cold pressor	3	−8, 18	IL-1β: **d_1 = −.06**	12
Shields et al., 2016 (48)	37	37	100%	not reported	Mixed	IL-1β, IL-6, IL-8,	passive drool	other (video of infant circumcision)	4	0, 10	IL-1β: **d_1 = .07** IL-6: **d_1 = .15** IL-8: **d_1 = .17**	8
Szabo et al., 2019 (40)	73	73	100%	70.42%	Healthy	IL-1β, IL-10	passive drool	TSST	10	0, 45	IL-1β: **d_1 = .29** IL-10: **d_1 = .50**	13
Szabo et al., 2016 (sample 1) (41)	46	46	71.74%	69.57%	Mixed	IL-10	passive drool	other (angry autobiographical memory)	10	−10, 50	IL-10: **d_1 = .38**	12
Szabo et al., 2016 (sample 2) (41)	45	45	68.89%	62.22%	Mixed	IL-10	passive drool	TSST	10	−10, 50	IL-10: **d_1 = .45**	12
Tell et al., 2018 (49)	30	30	100%	84%	Clinical (Breast Cancer)	IL-6	Salivette/ cotton	TSST	8	0, 23, 38, 53, 68	IL-6: **d_1 = .12**, d_2 = .05, d_3 = −.07, d_4 = −.11	9
Usui et al., 2011 (78)	10	10	100%	0%	Healthy	IL-1β, IL-6, TNF-α,	Salivette/ cotton	exercise	60	−10, 60, 120, 180	IL-1β: **d_1 = 10.67**, d_2 = 3.33, d_3 = 1.33 IL-6: **d_1 = 1.88**, d_2 = .42, d_3 = .50 TNF-α: **d_1 = 2.41**, d_2 = .88, d_3 = .40	10

Note. Original N's reflect the total sample size originally reported in the paper. Analytic N's reflect the total sample size after removing missing data or using all available data reported from authors, which were used in effect size calculations for the meta-analyses. Analytic N's are expressed as a range, as they varied slightly by biomarker and by time point.

*This study recruited overweight but subjectively healthy participants. Though originally coded as 'other,' it was coded as healthy for analysis.

†= effect sizes for these studies were aggregated due to multiple conditions or saliva collection methods. For results summary, d_1 = effect size (Cohen's d_{av}) from time 1 to time 2, d_2 = effect size from time 1 to time 3, d_3 = effect size from time 1 to time 4, d_4 = effect size from time 1 to time 5, d_5 = effect size from time 1 to time 6, d_6 = effect size from time 1 to time 7, d_7 = effect size from time 1 to time 8, d_8 = effect size from time 1 to time 9. **Bolded** effect sizes indicate the largest effect size, which were used in all meta-analyses (i.e., either the largest positive effect size, or the smallest negative effect size for studies with no positive effect sizes). CRP = C-reactive protein. INF-α = interferon alpha. INF-γ = interferon gamma. IL-1β = interleukin 1 beta. IL-2 = interleukin-2. IL-4 = interleukin-4. IL-5 = interleukin-5. IL-6 = interleukin-6. IL-8 = interleukin-8. IL-10 = interleukin-10. IL-12p70 = interleukin 12p70. IL-13 = interleukin 13. IL-17a = interleukin 17a. GM-CSF = granulocyte-macrophage colony-stimulating factor. TNF-α = Tumor necrosis factor alpha. TSST = Trier Social Stress Test. PASAT = Paced Auditory Serial Addition Task. Given that Acurio 2014 used a range of stressor lengths, we did not include this study in stressor length moderator analyses.

Table 2

Descriptives on Timing of Peak Effect Size and Stressor Length by Biomarker

	Peak Effect Size (in minutes from the start of the stressor)					Stressor Length (in minutes)	
	Mean	**SD**	**Median**	**Min**	**Max**	**Mean**	**SD**
IL-6	47.86	39.80	36.00	10.00	180.00	27.83	34.91
IL-1β	52.87	50.68	47.50	7.00	240.00	20.28	20.16
IL-10	98.68	62.61	70.00	45.00	240.00	36.41	32.78
CRP	40.51	20.86	56.20	15.00	60.00	34.37	15.22
TNF-α	66.52	63.18	50.00	23.66	240.00	34.55	33.19
IL-2	100.00	78.42	70.00	60.00	240.00	58.00	38.83
IL-4	100.00	78.42	70.00	60.00	240.00	58.00	38.83
IL-8	85.00	79.18	65.00	10.00	240.00	30.50	24.35
IL-12p70	30.00	0.00	30.00	30.00	30.00	25.00	0.00
INF-γ	65.00	5.77	65.00	60.00	70.00	42.50	20.21
IL-5	65.00	5.77	65.00	60.00	70.00	42.50	20.21

Note. IL-6 = interleukin-6, IL-1β = interleukin 1 beta, IL-10 = interleukin-10, CRP = C-reactive protein, TNF-α = Tumor necrosis factor alpha, IL-2 = interleukin-2, IL-4 = interleukin-4, IL-8 = interleukin-8. IL-12p70 = interleukin 12p70, INF-γ = interferon gamma, IL-5 = interleukin-5