Psychological stress reactivity and future health and disease outcomes: A systematic review of prospective evidence

Turner, A.I., Smyth, N., Hall, S.J., Torres, S.J., Hussein, M., Jayasinghe, S.U., Ball, K. and Clow, A.

Psychological stress reactivity and future health and disease outcomes: A systematic review of prospective evidence

Anne I Turner[a,*], Nina Smyth[b], Sarah J Hall[a], Susan J Torres[a], Mais Hussein[a], Sisitha U Jayasinghe[a,c], Kylie Ball[a], Angela J Clow[b]

[a]Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, Geelong, Victoria 3125, Australia

[b]Psychophysiology and Stress Research Group, Department of Psychology, University of Westminster, London, W1W 6UW, United Kingdom

[c]School of Health Sciences, College of Health and Medicine, University of Tasmania, Newnham, Tasmania 7250, Australia

Email address of each author:

anne.turner@deakin.edu.au, smythn@westminster.ac.uk, sarah.hall@deakin.edu.au, susan.torres@deakin.edu.au, mais.hussein18@gmail.com, sisitha.jayasinghe@utas.edu.au, kylie.ball@deakin.edu.au, clowa@wmin.ac.uk

For submission to: Psychoneuroendocrinology

*Corresponding Author:

Anne Isabella Turner, PhD

Institute for Physical Activity and Nutrition

School of Exercise and Nutrition Sciences

Deakin University

Melbourne Burwood Campus,

221 Burwood Highway, Burwood,

VIC 3125,

Australia

+61 3 92446950

anne.turner@deakin.edu.au

Abstract

Background

Acute psychological stress activates the sympatho-adrenal medullary (SAM) system and hypothalamo-pituitary adrenal (HPA) axis. The relevance of this stress reactivity to long-term health and disease outcomes is of great importance. We examined prospective studies in apparently healthy adults to test the hypothesis that the magnitude of the response to acute psychological stress in healthy adults is related to future health and disease outcomes.

Methods

We searched Medline Complete, PsycINFO, CINAHL Complete and Embase up to 15 Aug 2019. Included studies were peer-reviewed, English-language, prospective studies in apparently healthy adults. The exposure was acute psychological stress reactivity (SAM system or HPA axis) at baseline. The outcome was any health or disease outcome at follow-up after ≥ 1 year.

Results

We identified 1,719 papers through database searching and 1 additional paper through other sources. Forty-seven papers met our criteria including 32,866 participants (range 30 – 4100) with 1-23 years of follow-up. Overall, one third (32%; 83/263) of all reported findings were significant and two thirds (68%; 180/263) were null. With regard to the significant findings, both exaggerated (i.e. high) and blunted (i.e. low) stress reactivity of both the SAM system and the HPA axis at baseline were related to health and disease outcomes at follow-up. Exaggerated stress reactivity at baseline predicted an increase in risk factors for cardiovascular disease and decreased telomere length at follow-up. In contrast, blunted stress reactivity predicted future increased adiposity and obesity, more depression, anxiety and PTSD symptoms, greater illness frequency, musculoskeletal pain and regulatory T-Cell percentage, poorer cognitive ability, poorer self-reported health and physical disability and lower bone mass.

Conclusion

Exaggerated and blunted SAM system and HPA axis stress reactivity predicted distinct physical and mental health and disease outcomes over time. Results from prospective studies consistently indicate stress reactivity as a predictor for future health and disease outcomes. Dysregulation of stress reactivity may represent a mechanism by which psychological stress contributes to the development of future health and disease outcomes.

Keywords: acute stress, blood pressure, heart rate, epinephrine, norepinephrine, cortisol, sympatho-adrenal medullary system, hypothalamo-pituitary adrenal axis, health outcomes, disease outcomes

1. Introduction

When we are exposed to acute stress, two main physiological systems are activated. The nervous system responds immediately via activation of the sympatho-adrenal medullary (SAM) system with consequent release of the catecholamines, adrenaline and noradrenaline, which in turn stimulate increases in heart rate (HR) and blood pressure (BP) (Chrousos, 2009). Activation of the SAM system is also reflected in elevation of salivary alpha-amylase levels (Skoluda et al., 2017; Strahler et al., 2017). The endocrine system also responds to acute stress via activation of the hypothalamo-pituitary adrenal (HPA) axis with consequent release of cortisol from the adrenal cortex (Chrousos, 2009). Activation of both systems is integral to our response to stress in order to prepare the body to deal with the threat encountered and return the body to a steady state once the threat has passed. In contrast to physical stressors, responses to acute psychological stressors (to active-coping stressors) are considered "metabolically unjustified", since fuel stores within the body are mobilised for activity that does not eventuate (Obrist, 1976). These metabolically unjustified responses in particular may have consequences for pathological states (Obrist, 1976).

For decades, researchers and clinicians have been interested in whether the magnitude of the response to acute psychological stress is related to health and disease outcomes, with many researchers worldwide committing copious time and resources to measuring psychological stress reactivity to better characterise responses to stress. Indeed, a Web of Science search for the phrase "psychological stress reactivity" reveals a large increase in papers published in this field across the last two decades: 131 in 1998, 258 in 2008 and 430 in 2018. To provide context for the significance of findings from such studies (especially from cross-sectional studies, from which causation cannot be determined (Greenhalgh, 1997)), it is essential to understand how acute psychological stress reactivity is related to future physical and mental health and disease outcomes. Despite there being a broad base of research activity in this field, the relevance of psychological stress reactivity to long-term physical and mental health and disease outcomes remains an emergent and hotly studied topic (Brindle et al., 2016; Ronaldson et al., 2016; Zhu et al., 2016; Carroll et al., 2017; Steptoe et al., 2017).

As early as 1981, it was proposed that high cardiovascular responses during acute psychological stressors (active-coping stressors) might reflect a high degree of susceptibility to later hypertension (Light, 1981; Obrist, 1981). Much work has tested this "reactivity hypothesis" and reviews (Krantz and Manuck, 1984; Cinciripini, 1986; Light, 1987; Steptoe, 2000; Taylor et al., 2003; Treiber et al., 2003; Phillips and Hughes, 2011) and meta-analyses (Gasperin et al., 2009; Chida and Steptoe, 2010)

have been published showing support for this hypothesis. Nevertheless, the reactivity hypothesis, as originally stated, has a narrow focus on exaggerated cardiovascular reactivity and cardiovascular disease outcomes. Many questions remain unanswered and gaps remain in our understanding of this topic: for example, is parallel reactivity of non-cardiovascular measures important for health (e.g. catecholamines, salivary alpha-amylase, cortisol)? Is blunted reactivity also related to future physical and mental health and disease outcomes? Can health and disease outcomes other than cardiovascular disease also be related to earlier psychological stress reactivity (e.g. physical and mental health and disease outcomes)?

While much of the early work in this field focussed on exaggerated reactivity, more recent studies have also considered the role of blunted reactivity in future health and disease outcomes, including outcomes that extend beyond cardiovascular disease (Carroll et al., 2017). As evidence in this field accumulates, disparate findings are spread across the literature of many different disciplines and disease states, meaning that findings in some fields may go unnoticed in other fields despite their potential relevance. Previous systematic syntheses of this evidence have focussed solely on cardiovascular reactivity (Gasperin et al., 2009; Chida and Steptoe, 2010) or on findings from specific data sets (Carroll et al., 2017). With so many unanswered questions remaining, a broad systematic review of findings in this field is warranted.

We have taken a very broad approach to synthesising the current evidence in this field. Our overarching aim was to review prospective evidence examining how reactivity to psychological stress in healthy adults is related to future health and disease outcomes. We tested the hypothesis that the magnitude (both exaggerated and blunted) of the response to acute psychological stress (both SAM system and HPA axis) in healthy adults is related to future health and disease outcomes (both physical and mental). Since studies using prospective designs comprise a stronger form of evidence than cross-sectional studies (Greenhalgh, 1997), we chose to focus our review on prospective evidence.

2. Methods

2.1 Search Strategy

This review was undertaken in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement (Moher et al., 2009). The review protocol was registered with PROSPERO on 19 Dec 2017 (Registration ID CRD42017084161). The PICO framework, which describes participants, interventions, comparisons and outcomes was utilised to help develop the research question and search strategy (Moher et al., 2009). A systematic search was conducted by one author (AIT) in Medline Complete, PsycINFO and CINAHL Complete (all via Ebscohost) and Embase databases to locate peer-reviewed journal articles published up until 15 Aug 2019. Search terms included psychological stress AND (cortisol OR adrenaline OR noradrenaline OR salivary alpha-amylase OR HR OR BP) AND reactivity AND longitudinal and related terms. Full search strategies including all search terms are available in supplementary material. If studies from our primary search alerted us to other relevant studies, these were included if they also met our criteria.

2.2 Study inclusion and exclusion criteria

Included studies were conducted in apparently healthy adult populations. They were observational, prospective, population-based cohort studies. As indicated in Figure 1, included studies imposed an acute psychological stressor at baseline and measured reactivity of one or more of the following variables: systolic BP (SBP), diastolic BP (DBP), HR, adrenaline, noradrenaline, salivary alpha-amylase and cortisol. These studies then followed-up the same participants after one year or longer and measured any outcome related to physical or mental health or disease.

Studies were excluded if they were written in a language other than English, if they were non-peer reviewed publications (e.g. theses and conference abstracts) or if they were not original research articles (e.g. reviews, meta-analyses and protocol papers). Animal studies were excluded, as were studies conducted in children or adolescents (i.e. participants were <18 years old at baseline), pregnant women or in women who were peri-parturient or lactating. Studies were included if recruitment of apparently healthy individuals occurred via a workplace or particular occupation (e.g. public servants, military recruits), but not if participants were selected for inclusion because of a pre-existing health or disease condition (e.g. borderline hypertension, depression or individuals recovering from substance abuse). If studies recruited using a population-based cohort approach and the sample included some individuals with a pre-existing condition at baseline (e.g. hypertension), these studies were included as long as there was no oversampling or special selection

for individuals with the condition and those present were there because of the natural presence of that condition in the population. Studies investigating psychological stress reactivity were included, while those investigating only physical stress reactivity were excluded (e.g. cold pressor test or handgrip test). Studies were excluded if reactivity to acute psychological stress at baseline was not considered alone, but only in combination with other variables (e.g. in combination with job demands or physical stressors). This review focused on SAM system (SBP, DBP, HR, adrenaline, noradrenaline and/or salivary alpha-amylase) and HPA axis (cortisol) responses to stress. Studies investigating only reactivity of other factors released during stress (e.g. cytokines) were beyond the scope of this review and were excluded.

2.3 Screening, data extraction and assessment of study quality and risk of bias

Search results were first imported into Endnote (X8, Clarivate Analytics, Philadelphia, USA), where duplicates were removed. To identify studies for inclusion, title and abstract screening and full-text screening were undertaken independently by two authors (AIT and SJH) using the Rayyan software platform (Ouzzani et al., 2016). Discrepancies were resolved by discussion and consensus. If discrepancies had remained, input from a third party would have been sought, but this was not necessary. For included studies, data extraction and assessment of quality and bias were undertaken independently by two parties (AIT and NS/AJC) after first pilot testing our draft data extraction spreadsheet using five studies selected from our included set. Discrepancies were resolved by discussion and consensus. If discrepancies had remained, input from a third party would have been sought, but this was not necessary.

For each study, we extracted study author names, year of publication, country in which the study took place, number and sex of participants included in prospective analyses, age at baseline, name of cohort study (if relevant), type of acute psychological stressor imposed at baseline, stress reactivity variables measured at baseline (including sample times and method of reactivity calculation for cortisol), length of follow-up, physical and/or mental health and/or disease variables measured at follow-up and reported study findings (significant and null) relevant to the objectives of our review. Where studies included findings not relevant to the objectives of our review (e.g. cross-sectional analyses, physical stressors, cytokine reactivity, stress recovery findings), these data were not extracted.

Quality and risk of bias for each study were assessed by one author (AIT) using a modified version of the Cochrane Tool to Assess Risk of Bias in Cohort Studies (Cochrane, London, UK). Our assessment utilised the following questions: i) Can we be confident in the assessment of exposure? ii) Can we be confident in the assessment of outcome? iii) Can we be confident that the outcome of interest was

not present at start of study? iv) Can we be confident in the assessment of the presence or absence of relevant confounders? v) Was the follow up of cohorts adequate? Answer options were: “Definitely yes”, “Probably yes”, “Probably no” and “Definitely no”, which were allocated numerical values of 4, 3, 2 and 1, respectively. One further question was asked: vi) Was there evidence of any other sources of bias in reporting? Answer options were: “no”, “possibly” or “yes”.

2.4 Synthesis of results

Synthesis of study findings was achieved by calculating the total number of significant and null findings and by summarising and drawing together the main significant findings of the studies in a tabular form under the headings exaggerated reactivity and blunted reactivity for each of the SAM system and the HPA axis. Study quality and risk of bias were considered at an individual study level and an overall level in the appraisal of study outcomes. Due to the vast number of different exposure variables (SBP, DBP, HR, adrenaline, noradrenaline, salivary alpha-amylase and/or cortisol) and outcome variables (any measure of health or disease outcome was acceptable), meta-analysis was not appropriate and would not be meaningful.

3. Results

3.1 Study selection

Our primary searches yielded 1,719 studies and we identified one additional study (Zhu et al., 2016) via our included studies (Figure 2). The paper by Zhu and colleagues (2016) was not identified in our primary search as the search terms “prospective”, “longitudinal” and “follow-up” did not appear in the title or abstract, nor was the study categorised under database assigned subject headings such as “Prospective Studies”, "Longitudinal Studies" or "Follow-Up Studies". After removal of duplicates (n = 565), 1,155 studies were screened by title and abstract and full-text articles were retrieved and screened for 80 studies. Thirty-three full-text articles did not meet our inclusion and/or exclusion criteria (see Figure 2 for details) and were excluded leaving 47 studies that were included in our systematic review.

3.2 Characteristics of included studies including assessment of quality and risk of bias

Table 1 shows the study details, study characteristics, significant and null findings and assessment of quality and risk of bias for each of the 47 included studies. Prospective studies in this field have been published over the last 25 years (1992 - 2017) at an average rate of 1.9 per year. The first SAM system study was published in 1992, while the first HPA axis study was published in 2012. Thirty-nine studies have investigated reactivity of the SAM system alone, six have investigated reactivity of the HPA axis alone and two have investigated reactivity of both systems within the same study. The total number of participants across the 47 included studies was 32,866 and the mean (±SD) number per study was 699 ± 921 (range 30 – 4100; median = 411). The mean (±SD) duration of follow-up was 6.4 ± 4.9 years (range 1 - 23 years; median = 5.0 years). Many large cohort studies are represented including the Whitehall II, CARDIA, West of Scotland Twenty-07 and Dutch Famine Birth Cohort studies (Table 1). Commonly studied psychological stressors have been used including mental arithmetic, video game, mirror image star-tracing task, structured interview, Raven’s matrices, anticipation of exercise, Stroop colour–word interference task, speech tasks and Trier Social Stress Test (TSST). A range of reactivity variables have been measured including SBP, DBP, mean arterial pressure (MAP), HR, adrenaline and noradrenaline from the SAM system and salivary cortisol and plasma cortisol from the HPA axis. None of the 47 included studies measured salivary alpha-amylase reactivity. Outcome measures have been many and varied including resting and ambulatory SBP and DBP, hypertensive status, carotid intima-media thickness (IMT), coronary artery calcification (CAC), body mass index (BMI), waist circumference, obesity status, telomere length, cognitive ability,

symptomology for depression, anxiety and post-traumatic stress disorder (PTSD), self-reported health, self-reported physical disability, musculoskeletal pain, regulatory T-Cell percentage and bone mineral content and density.

Within individual studies, quality and risk of bias is variable with the full range of assessment response options (1-4) having been used (Table 1). For this body of work as a whole (i.e. across studies), study quality is generally high and risk of bias is generally low (Figure 3). The percentage of studies for which the answer was "Definitely yes" or "Probably yes" was 94% for "Can we be confident in the assessment of exposure?"; 94% for "Can we be confident in the assessment of outcome?"; 87% for "Can we be confident that the outcome of interest was not present at start of study?"; 94% for "Can we be confident in the assessment of the presence or absence of relevant confounders?"; and 87% for "Was the follow up of cohorts adequate?" In answer to the question, "Was there evidence of any other sources of bias in reporting?", 45 studies were deemed "no", 2 studies (Steptoe and Wardle, 2005; Senan and Petrosyan, 2014) were deemed "possibly" and none were deemed "yes". While the stated objective of the study by Steptoe and Wardle (2005) was to investigate the influence of recovery (rather than reactivity) on adiposity measures over 3 years, the authors state that the extent to which reactivity predicts changes in BMI and waist-hip ratio (WHR) was also assessed. Therefore, the omission of results for the influence of reactivity on WHR change (see Table 1) may reflect reporting bias, but may instead reflect that these results were not deemed important to the stated objectives of the study. The paper by Senan and Petrosyan (2014) provides insufficient details to assess if there is reporting bias. For example, the paper does not include a description of the statistical analyses conducted.

3.3 Synthesis of results

Of the 47 studies included in this review, 39 (83%) reported one or more significant findings and 39 (83%) reported one or more null findings (Table 1). Overall, one third (32%; 83/263) of reported findings were significant and two thirds (68%; 180/263) were null (Table 1). Significant findings are synthesised in Table 2. Both exaggerated and blunted reactivity of both the SAM system and HPA axis were related to future physical and mental health and disease outcomes. Overall, the pattern of significant vs null findings was not influenced by the type of stress task utilised, the duration of follow-up period or the timing of samples or method of calculating reactivity for cortisol.

Exaggerated cardiovascular reactivity (mostly exaggerated SBP and DBP reactivity) was related to greater future resting and ambulatory SBP and DBP, risk of hypertension, earlier onset of hypertension, greater progression of IMT, odds of having CAC and increased risk of sudden coronary

death and death from any cause (Table 2). Exaggerated noradrenaline reactivity at baseline was related to greater waist circumference at follow-up.

Exaggerated salivary cortisol reactivity was related to greater risk of hypertension and greater CAC progression (Table 2). Beyond cardiovascular outcomes, findings for exaggerated salivary cortisol reactivity at baseline also showed a relationship to shorter leukocyte telomeres at follow-up.

Blunted cardiovascular reactivity (mostly blunted HR reactivity) was related to greater illness frequency, greater increases in resting DBP, greater odds of having CAC in blacks (but not whites), greater carotid IMT, greater incidence of obesity, poorer self-reported health and greater progression of self-reported physical disability, more symptoms for depression and anxiety and reduced cognitive ability (Table 2). Blunted adrenaline reactivity was related to greater BMI, waist circumference and triceps skin thickness.

Blunted plasma cortisol reactivity was prospectively related to greater musculoskeletal pain in pain-sensitive females (but not males) (Table 2). Blunted salivary cortisol was related to more PTSD symptomology after exposure to new-onset traumatic events, greater regulatory T-Cell percentage (which in turn, was associated with poorer physical and mental health and more depressive symptomatology) and lower bone mineral content and bone mineral density in males (but not females; Table 2).

With regards to null findings, eight studies (17%) did not find any significant link between reactivity at baseline and health and disease outcomes at follow-up (Table 1). While there were no obvious differences overall in study quality or risk of bias across these vs studies with significant findings, each of these studies had individual limitations that might help explain null findings. Since Brody et al. (1996) measured MAP reactivity at baseline rather than SBP and DBP separately, it is not known if significant findings may have been present for either SBP or DBP alone. Fauvel et al. (2003) split participants into high-SBP-reactivity and non-high-SBP-reactivity groups with the cut-off between groups based on criteria unrelated to stress reactivity (based on the level of perceived job strain). Detailed results for linear multivariate analyses were not presented and, despite measuring DBP and HR reactivity at baseline, results for these variables are not reported (Fauvel et al., 2003). In the study by Steptoe and Marmot (2005), significant findings for reactivity did not survive adjustment for post-stress recovery, which was an absolute value, measured 40-45 min after the stressor. Since there was no significant change in BMI across the 3 year study by Steptoe and Wardle (2005), it is not surprising that reactivity did not predict BMI change, and as mentioned earlier, despite being measured, results for reactivity and change in WHR are missing from this paper. Despite rigorous methods in the study by Stewart et al. (2006), there was an unexpected but significant decrease in

resting SBP and DBP from baseline to follow-up (3.2 years). In light of this finding, it is not surprising that reactivity did not predict change in resting SBP or DBP. Flaa et al. (2008a) found that adrenaline and noradrenaline reactivity did not predict fasting glucose or HOMA-IR after 18 years of follow up, but it is not known if significant findings for SBP, DBP or HR reactivity may have been present, as these analyses were not conducted/reported (despite being measured at baseline (Flaa et al., 2008b)). While Shaffer et al. (2012) found that SBP and DBP reactivity did not predict risk of incident CVD events during 10 years of follow-up, this study does not appear to have achieved reliable and consistent stress reactivity at baseline. Mean SBP reactivity (mean ± SD = 1.1 ± 11.0 mmHg) and DBP reactivity (mean ± SD = 2.2 ± 7.4 mmHg) were close to 0 mmHg with some individuals having large negative SBP reactivity (range = -39 to 48 mmHg) and DBP reactivity (range = -22 to 27 mmHg) in response to the stressor (Shaffer et al., 2012). Finally, while Gentile et al. (2015) found that SBP, DBP and HR reactivity did not predict metabolic burden (measured as the number of metabolic parameters for which participants were in the highest quartile; lowest for high-density lipoprotein cholesterol) or incident metabolic syndrome, it is not clear if blood sampling (for metabolic makers) may have interfered with reactivity measures. The timing and method of blood sample collection is not reported in relation to the timing of measurement of SBP, DBP and HR response to psychological stressors.

It must also be noted that studies reporting significant findings often also reported a number of null findings (Table 1). Indeed, overall, more null findings (180/263; 68%) than significant findings (83/263; 32%) were reported. As an example, while SBP reactivity at baseline predicted resting SBP at follow-up (Carroll et al., 2001), DBP reactivity did not predict resting DBP and neither SBP nor DBP predicted hypertensive status at follow-up. In another example (Phillips et al., 2011a), while blunted HR reactivity at baseline predicted greater deterioration of physical function over the 5 years of follow-up, SBP and DBP reactivity did not predict change in disability score. In one final example (this time from the HPA axis), blunted salivary cortisol reactivity at baseline predicted lower BMD and BMC at follow-up in males, but not in females, and these findings did not extend to total plasma cortisol reactivity, which did not predict BMD or BMC at follow-up (Zhu et al., 2016).

Effect sizes for significant findings were typically modest and consistent across different studies. HR reactivity to mental arithmetic accounted independently for 4% of the variance in resting SBP at follow-up in the study by Stewart and France (2001). In the study by Carroll et al. (2003), SBP and DBP reactivity at baseline predicted 5-year upward drift in resting SBP and DBP, respectively, accounting for 3.6% and 2.9% of variance, respectively, beyond traditional risk factors. Finally, 4.5% of the variance in PTSD symptomatology in soldiers who had experienced new-onset traumatic events was explained by salivary cortisol reactivity at baseline, with lower (blunted) reactivity

predicting greater increase in PTSD symptomatology (Steudte-Schmiedgen et al., 2015). Effect sizes were consistent across studies using different types of stress tasks, different reactivity variables and with varying outcomes investigated. There were no health or disease outcomes that were more prominent than others, based on effect sizes.

4. Discussion

This systematic review presents a consistent line of prospective evidence that both exaggerated and blunted responses of both the SAM system and HPA axis to acute psychological stress in healthy adults are related to future physical and mental health and disease outcomes. One third of all reported findings linking stress reactivity at baseline to health and disease outcomes at follow-up were significant while two thirds were null. In relation to the significant findings, the reactivity hypothesis has proven to apply in a much broader context than originally proposed. The original reactivity hypothesis had a narrow focus on the relationship between exaggerated cardiovascular reactivity and future cardiovascular disease outcomes (Light, 1981; Obrist, 1981). Our systematic review provides substrate for a "bidirectional multi-system reactivity hypothesis" that the magnitude (both exaggerated and blunted) of the response to acute psychological stress (both SAM system and HPA axis) in healthy adults is related to future health and disease outcomes (both physical and mental). In relation to the null findings, a simple count of findings may provide a useful overview, but may also give an overly simplistic assessment of this field. In contrast, a detailed assessment of combinations of outcomes within and between studies may reveal a richness of information about underlying mechanisms and patterns of response that may be important in determining the relationship between reactivity to psychological stress and future health and disease outcomes. The specific outcomes reported in this field to date fall into several categories, which are considered in the following subsections (4.1 - 4.4).

4.1 Cardiovascular disease risk factor outcomes

In keeping with the original reactivity hypothesis, there is no surprise that individuals with elevated cardiovascular reactivity at baseline (SBP, DBP and/or HR) had more risk factors for cardiovascular disease at follow-up including greater resting and ambulatory SBP and DBP, more incident hypertension, earlier onset of hypertension, greater progression of IMT, greater risk of having CAC and increased risk for sudden coronary death (Table 2). Interestingly, three studies also reported that individuals with blunted HR reactivity had more risk factors for cardiovascular disease at follow-up including greater increases in resting DBP (Brody and Rau, 1994), greater odds of having CAC in blacks (but not whites) (Matthews et al., 2006) and greater carotid IMT (Heponiemi et al., 2007). These findings expand the original reactivity hypothesis to implicate blunted as well as exaggerated reactivity in the pathogenesis of cardiovascular disease, thus, supporting an inverted-U model of the reactivity hypothesis in which exaggerated and blunted responses may be detrimental but intermediate responses may be the most adaptive and resilient responses (Carroll et al., 2017). Thus, the reactivity hypothesis has proven its utility in a bidirectional way, incorporating both exaggerated

and blunted stress reactivity in relation to future health and disease outcomes. It is noteworthy that HR was the reactivity variable in each case where blunted cardiovascular reactivity was related to future cardiovascular disease risk factors. These findings are in keeping with results of a cluster analysis by Brindle et al. (2016), whereby the cluster with the greatest risk for future hypertension was one in which SBP and DBP reactivity was exaggerated, but HR reactivity was relatively small. Since MAP is determined by changes in cardiac output and/or total peripheral resistance (i.e. MAP = cardiac output X total peripheral resistance), this may suggest maladaptation when MAP is elevated primarily via peripheral vasoconstriction rather than via elevation of cardiac output (Brindle et al., 2016). Indeed, CAC was predicted by both exaggerated SBP reactivity and blunted HR reactivity (in blacks) within the same study (Matthews et al., 2006). The most adaptive responses may involve the elevation of MAP primarily via increases in cardiac output. These findings suggest that the combination or pattern of response variables may be important in relation to future health and disease outcomes.

Keeping, for now, with cardiovascular disease risk factor outcomes, exaggerated cortisol reactivity at baseline was also implicated. Findings that exaggerated salivary cortisol reactivity predicted greater risk of hypertension (Hamer and Steptoe, 2012) and greater CAC progression (Hamer et al., 2012) effectively extend the reactivity hypothesis beyond just the SAM system to include the HPA axis, as well. These findings provide support for a multi-system reactivity hypothesis by providing evidence that stress reactivity of multiple stress response systems (SAM system and HPA axis) is related to future health and disease outcomes. There are known to be bidirectional stimulatory connections between the HPA axis control centre in the hypothalamus and the SAM system control centre in the brainstem, such that activation of either one of these systems results in activation of the other (Chrousos, 2009). Consequently, it is tempting to speculate that individuals with exaggerated SAM system reactivity may be those who also display exaggerated HPA axis reactivity. However, the findings of Hamer et al. (2012) do not support this supposition. While salivary cortisol reactivity at baseline predicted CAC progression, SBP and DBP reactivity at baseline in the same participants did not predict CAC progression at follow-up. No other prospective studies have measured both cortisol reactivity and cardiovascular reactivity together in the same study in relation to cardiovascular disease outcomes.

4.2 Adiposity outcomes

Findings for adiposity outcomes are interesting. The greatest weight of evidence shows that blunted reactivity is related to accumulation of adipose tissue over time. For example, two studies found that, after adjusting for a range of possible confounders including baseline obesity status, blunted

HR reactivity predicted greater future likelihood of being obese (Carroll et al., 2008; Phillips et al., 2012). Furthermore, blunted adrenaline reactivity was related to greater future BMI, waist circumference and triceps skinfold thickness (Flaa et al., 2008c). Nevertheless, further findings from this latter study showed that exaggerated noradrenaline reactivity was also related to greater waist circumference in the same participants (Flaa et al., 2008c). In other words, greater waist circumference at follow-up was predicted by both blunted adrenaline reactivity and exaggerated noradrenaline reactivity in the same participants. Since adrenaline has a slightly greater effect on the heart (via beta-adrenergic receptors) and noradrenaline on peripheral vasoconstriction (via alpha-adrenergic receptors) (Brunton et al., 2018), this profile of reactivity responses aligns with the exaggerated blood pressure and blunted heart rate profile described earlier. Thus, these findings provide further support for the notion that the pattern of response is important in relation to future health and disease outcomes. These findings also provide further support for an inverted-U model of the reactivity hypothesis, in which exaggerated and blunted responses may be detrimental, but intermediate responses may be the healthiest responses (Carroll et al., 2017). In the only study to measure cortisol reactivity in relation to future adiposity (Phillips et al., 2012), no relationship to future adiposity was found.

4.3 Self-report, mental health and cognitive outcomes

Self-report, mental health and cognitive outcomes were always predicted by blunted as opposed to exaggerated reactivity. Blunted cardiovascular reactivity predicted poorer outcomes for self-reported illness frequency (Lawler and Schmied, 1992), self-reported health (Phillips et al., 2009), progression of self-reported physical disability (Phillips et al., 2011a), more depression symptoms (Phillips et al., 2011b), more anxiety symptoms (Yuenyongchaiwat and Sheffield, 2017) and poorer cognitive ability (Ginty et al., 2011; Yuichiro et al., 2016). Blunted HPA axis reactivity was also implicated, with blunted salivary cortisol reactivity predicting more PTSD symptomology in male soldiers following exposure to new-onset traumatic events during military deployment to Afghanistan (Steudte-Schmiedgen et al., 2015). Collectively, these behavioural, mental health and cognitive findings provide support for a compelling argument proposed by Carroll et al. (2017) that blunted reactivity of the stress systems may reflect dysregulation of the motivational systems within the brain.

4.4 Other outcomes

Several other disparate outcomes were also related to baseline levels of HPA axis reactivity (both exaggerated and blunted). Exaggerated cortisol reactivity predicted telomere attrition (Steptoe et al., 2017), whereas blunted cortisol reactivity predicted musculoskeletal pain in pain-sensitive

females (Paananen et al., 2015), regulatory T-cell percentage, which in turn was related to poorer physical and mental health and more depressive symptomatology (Ronaldson et al., 2016), and lower bone mineral content and density in males (Zhu et al., 2016). Interestingly, in the study by Zhu et al. (2016), salivary cortisol but not total plasma cortisol predicted lower BMC and BMD in males. This discrepancy may have occurred because the study was better powered for salivary cortisol measures, with more of the participants in the study providing samples for salivary cortisol measurement (n=801; 390 females and 411 males) compared with plasma cortisol measurement (n=648; 291 females and 357 males). Alternatively, the discrepancy between salivary cortisol and total plasma cortisol findings may reflect underlying differences in the sensitivity of these measures in detecting dynamic activity of the HPA axis (Zhu et al., 2016). Since salivary cortisol represents the free fraction of circulating cortisol (i.e. that fraction which is not attached to a binding protein), it is thought to be a more sensitive measure of dynamic HPA axis activity (Kirschbaum and Hellhammer, 1994; Gozansky et al., 2005). The major plasma binding protein for cortisol (corticosteroid-binding globulin; CBG) is known to vary widely within and between individuals and has been shown to complicate the interpretation of dynamic assessment of HPA axis activity as measured by total plasma cortisol (Dhillo et al., 2002). Salivary cortisol collection has the added advantage of being minimally invasive compared with plasma cortisol collection. Nevertheless, despite salivary cortisol being a more sensitive measure of dynamic HPA axis activity, Paananen et al. (2015) found total plasma cortisol to be a significant predictor of musculoskeletal pain in pain-sensitive females, perhaps reflecting a greater effect size for this outcome. Salivary cortisol was not measured in this study (Paananen et al., 2015). None of these disparate studies of cortisol reactivity included any measures of SAM system reactivity.

4.5 Implications of significant findings

Collectively, these significant findings support a “bidirectional multi-system reactivity hypothesis”. We have shown that a diverse range of physical and mental health outcomes are predicted by both exaggerated and blunted responses of the SAM system and the HPA axis to acute psychological stress. Healthy adaptive stress responses appear to be intermediate or “Goldilocks” stress responses: Not too large and not too small (Figure 4). While we cannot conclude causation from prospective cohort studies, the temporal sequence of these studies, with exposure preceding outcome, provides a stronger form of evidence towards potential causality compared with cross-sectional studies in which the exposure and outcome coincide (Greenhalgh, 1997). Indeed, many authors have speculated that stress reactivity may be a mechanism or possible causal factor by which psychological stress contributes to health and disease outcomes (Flaa et al., 2008b; Hamer et

al., 2012; Hamer and Steptoe, 2012; Steptoe et al., 2016; Yuenyongchaiwat, 2017). Consequently, the role of stress reactivity in future health and disease outcomes deserves further consideration.

The pattern of response may be an important factor in the link between stress reactivity and future health and disease outcomes. For the SAM system, exaggerated blood pressure reactivity combined with blunted cardiac reactivity appears to be emerging as a maladaptive response. While the combination of responses between the SAM system and HPA axis may also be important, relatively fewer studies have considered the HPA axis and only two have considered reactivity of both systems in the same participants within the same study (Hamer et al., 2012; Phillips et al., 2012). Nevertheless, there appears to be some consistency in outcomes across these two systems as both exaggerated SAM system and exaggerated HPA axis reactivity predicted cardiovascular risk factor outcomes and both blunted SAM system reactivity and blunted HPA axis reactivity predicted mental health symptomology (Table 2). Further consideration of the role of response patterns is warranted.

Given the findings of consistent links between stress reactivity and future health and disease outcomes, it is tempting to contemplate the utility of stress reactivity as a clinical screening tool for early diagnosis of exaggerated or blunted reactivity to facilitate identification of individuals for targeted intervention. Nevertheless, this eventuality seems unlikely. A barrier to the use of stress reactivity as a clinical screening tool is our lack of reference values for the magnitude of stress reactivity that is linked to future adverse health and disease outcomes. Furthermore, the development of reference values for stress reactivity would seem difficult at best given the vast array of methodological differences utilised in this body of research. Research conducted to date has used many different stressors, many different reactivity variables (SBP, DBP, MAP, HR, adrenaline, noradrenaline and cortisol) and many different sampling protocols with no standardisation of the number or timing of measures taken. Different ways of calculating reactivity have also been used. For cortisol, best-practice recommendations for calculating reactivity are available (Miller et al., 2018), but consideration may also need to be given to the ultradian rhythmicity of cortisol (Henley et al., 2009; Lightman and Conway-Campbell, 2010) and the debate around the classification of responders versus non-responders (Miller et al., 2013). Consensus agreement on best-practice methodologies for measuring stress reactivity would be needed, as well as an increase in the number of studies that measure multiple stress systems in the same participants within the same studies.

Effect sizes found in this review have generally been modest and consistent across studies indicating that stress reactivity offers relatively low levels of independent prediction of health and disease outcomes beyond other (traditional) risk factors and covariates. Nevertheless, reactivity to

psychological stress may represent a modifiable risk factor, whereas some other traditional health and disease risk factors and covariates are not modifiable (e.g. age and sex). If reactivity to psychological stress is a causal component in the aetiology of disease, these findings support a role for programs aimed at reducing stress reactivity such as stress management in daily life (e.g. physical activity, mindfulness and/or nature relatedness programs). Since the link between stress reactivity at baseline and health and disease outcomes at follow-up is similar across different types of outcomes, stress management programs may have equal benefits across multiple dimensions of health and disease.

4.6 Implications of null findings

In relation to the null findings, it is interesting to note that overall, null findings outnumbered significant findings two to one. While this summary statistic helps to give an overall indication of the state of this field, a simple count of findings may also give an overly simplistic assessment. In many cases, null findings showed a trend towards being significant (P = 0.05 - 0.1) and/or represented significant findings that did not survive adjustment for covariates. A detailed assessment of combinations of outcomes within and between studies instead reveals a richness of information about underlying mechanisms and patterns of response that may be important in determining the relationship between reactivity to psychological stress and future health and disease outcomes. For example, that SBP but not DBP or HR reactivity predicted IMT (Jennings et al., 2004) may indicate nuances of underlying physiological mechanisms at play rather than simply a lack of utility of DBP and HR reactivity as predictive factors. Similarly, that blunted DBP and HR but not SBP reactivity predicted poorer self-reported health at follow-up (Phillips et al., 2009) might point towards mechanisms at play rather than a simple lack of utility of SBP as a predictive factor for this outcome. For cortisol, the prediction of BMD and BMC by blunted salivary but not plasma cortisol reactivity (Zhu et al., 2016) may be an informative methodological consideration rather than being a useful contribution to an overall summary statistic.

4.7 Strengths and limitations

Given the maturity of this field, we have been able to limit our review to prospective evidence. This is a strength of our review. Prospective studies provide a stronger form of evidence than cross-sectional studies because of the temporal sequence between exposure and outcome (Greenhalgh, 1997) (i.e. reactivity is measured at baseline, before any health or disease outcomes emerge and are measured at follow-up). In keeping with best practice, most studies included in this review adjusted for baseline levels of outcome variables to ensure that the outcome variable of interest was not present at the start of the study. Indeed, for 87% of studies, we answered “Definitely yes” or

"Probably yes" for "Can we be confident that the outcome of interest was not present at start of study?"

Even though our search of the literature and our synthesis of findings have been systematic, our conclusions may be limited if publication bias exists in the field. Researchers and journals may be more likely to publish studies reporting significant findings compared with null findings. Indeed, some included studies have reported investigating variables for which no findings are reported (Matthews et al., 1993; Fauvel et al., 2003; Steptoe and Wardle, 2005; Deter et al., 2006; Carroll et al., 2008; Phillips et al., 2012; Yuenyongchaiwat et al., 2015; Yuenyongchaiwat, 2017) (see Table 1). It is not clear why such findings were omitted. Perhaps *a priori* decisions were made about which exposure/outcome combinations were most important to investigate. Alternatively, if findings were null, perhaps there was a perception that such findings were less important to report or perhaps journal space limitations precluded reporting of null findings.

The consideration of and adjustment for potential confounding factors was undertaken in most included studies. Indeed, for 94% of studies, we answered "Definitely yes" or "Probably yes" for "Can we be confident in the assessment of the presence or absence of relevant confounders?" Nevertheless, it is possible that there is residual confounding by unmeasured confounding variables in this body of work. A potential confounding variable that deserves investigation in future studies is early life adversity (Bunea et al., 2017; Lovallo et al., 2017), which may influence both our exposure variable (stress reactivity) and our outcome measures (health and disease outcomes).

4.8 Future directions

There are a number of key future directions that would be beneficial for this field going forward:

Firstly, given the bidirectional stimulatory connections between the SAM system and HPA axis (Chrousos, 2009) and the emerging importance of the pattern of reactivity response variables, it would be beneficial for further studies to consider reactivity of multiple stress systems within the same study. Only two studies in the current review measured both SAM system and HPA axis reactivity within the same study (Hamer et al., 2012; Phillips et al., 2012). Consideration of the patterns of response of reactivity variables both within the SAM system and across this and the HPA axis may reveal important further insights regarding the role of stress reactivity in future health and disease outcomes.

Secondly, while it was beyond the scope of the current review, reactivity of the immune system to acute psychological stress may represent an important additional pathway for investigation. Systemic inflammatory markers are known to increase in response to acute psychological stress

(Rohleder, 2014), with several studies implicating reactivity of inflammatory markers in future health and disease outcomes (Brydon and Steptoe, 2005; Ronaldson et al., 2016; Steptoe et al., 2016).

Thirdly, in addition to patterns of individual response variables, consideration of the integrated response of the stress pathways may also be enlightening. Some work has considered the integrated response of the SAM system and the HPA axis by measuring the ratio of the response of these two systems (Ali and Pruessner, 2012) and this approach may prove valuable in advancing this field in future studies.

Fourthly, to broaden our understanding of the range of outcome measures linked to stress reactivity, opportunities may exist for further hypothesis testing using existing data sets where reactivity has been measured at baseline. Indeed, there may be great scope to expand our understanding in this field further using existing data sets.

Lastly, in light of the findings of this review, further systematic syntheses would also help advance our understanding of underlying mechanisms in this field:

- It would be beneficial to consider the *characteristics of individuals that determine the magnitude of their response to acute psychological stress.* A systematic synthesis of existing research could consider factors such as genetic variation (e.g. Wust et al. (2004)), early life adversity (e.g. Bunea et al. (2017)) and levels of adiposity (e.g. Jones et al. (2012)) and physical activity (e.g. Mucke et al. (2018)).
- Another natural extension of this research is to determine *whether the magnitude of the response to acute psychological stress can be modified using interventions* aimed at improving certain aspects of lifestyle behaviours such as mindfulness interventions (e.g. Lindsay et al. (2018)), dietary interventions (e.g. West et al. (2012)) and exercise interventions (e.g. Throne et al. (2000)). A systematic synthesis of existing research on this topic would be valuable.

5. Conclusions

This systematic review provides substrate for a “bidirectional multi-system reactivity hypothesis” that the magnitude (both exaggerated and blunted) of the response to acute psychological stress (both SAM system and HPA axis) in healthy adults is related to future health and disease outcomes (both physical and mental). Our findings provide support for an inverted-U model of the reactivity hypothesis in which exaggerated and blunted responses may be detrimental, but intermediate responses, “Goldilocks” responses, may be the most adaptive and resilient responses. Our findings

also provide support for the notion that the pattern of individual response variables may be important to the link with future health and disease outcomes. In future research, this field would benefit from considering multiple stress system pathways and response patterns in relation to stress reactivity and additional health and disease outcomes. It will also be beneficial to consider factors that determine the magnitude of stress reactivity and whether interventions aimed at improving lifestyle behaviours can improve stress reactivity. If stress reactivity is a mechanism through which psychological stress contributes to the development of future health and disease outcomes, these findings provide support for the benefits of stress management programs.

Declaration of interests

The authors have no conflicts of interest to declare.

CRediT author statement

Anne I Turner: Conceptualization, Methodology, Investigation, Formal Analysis, Visualization, Writing – Original Draft, Writing – Review & Editing. **Nina Smyth:** Conceptualization, Formal Analysis, Writing – Review & Editing. **Sarah J Hall:** Formal Analysis, Writing – Review & Editing. **Susan J Torres:** Conceptualization, Writing – Review & Editing. **Mais Hussein:** Conceptualization, Writing – Original Draft, Writing – Review & Editing. **Sisitha U Jayasinghe:** Conceptualization, Writing – Review & Editing. **Kylie Ball:** Conceptualization, Writing – Review & Editing. **Angela J Clow:** Conceptualization, Methodology, Formal Analysis, Writing – Review & Editing.

Acknowledgements

We thank Deakin University liaison librarian, Rachel West, for her guidance with the design and implementation of the systematic search of the literature.

Funding

AIT received Academic Studies Program funding from Deakin University. The funding body had no involvement in study design, collection, analysis or interpretation of data, in writing the report or in the decision to submit the article for publication.

Supplementary material

Full search strategies are available as supplementary material

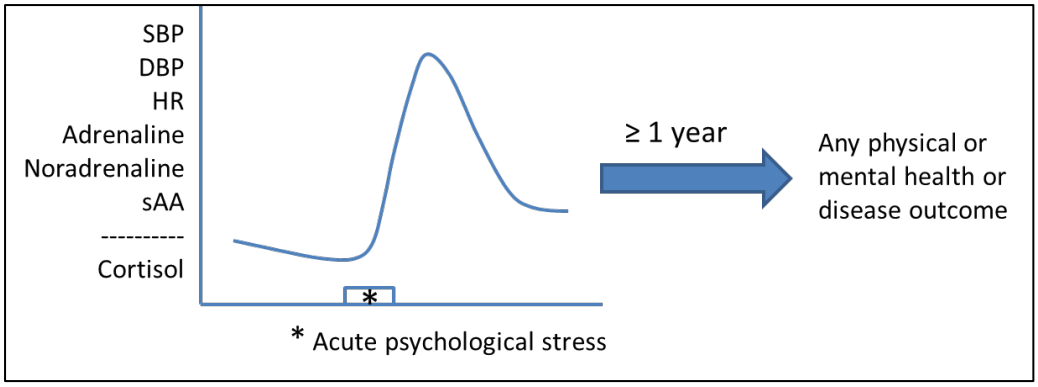


Figure 1. Schematic representation of systematic review criteria illustrating a typical reactivity response. Included studies imposed an acute psychological stressor at baseline and measured reactivity of one or more of systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate (HR), adrenaline, noradrenaline, salivary alpha-amylase (sAA) and/or cortisol. The same participants were then followed-up after one year or longer for the measurement of any physical or mental health or disease outcome.

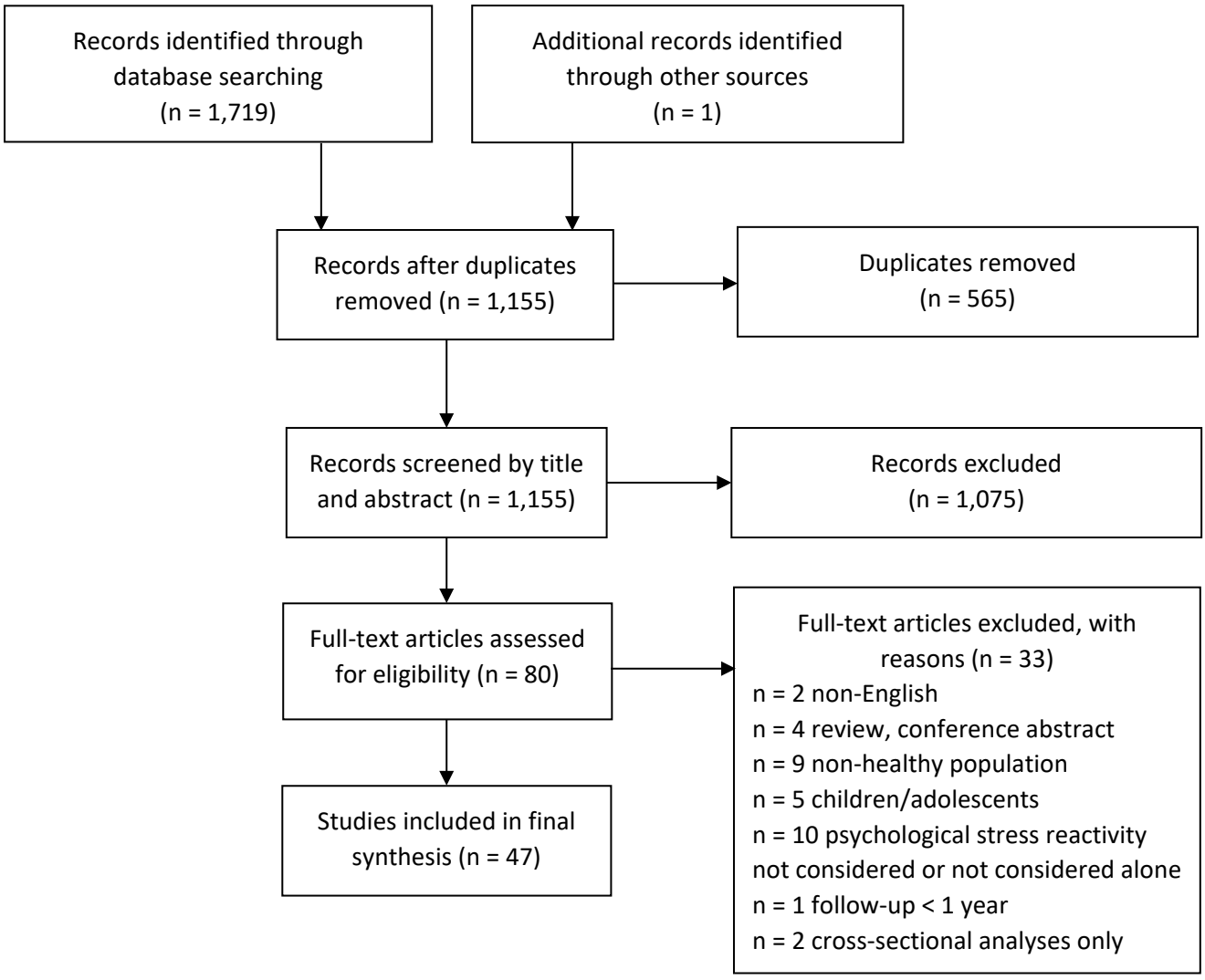


Figure 2. PRISMA flow diagram of studies identified, screened, excluded and included in this systematic review.

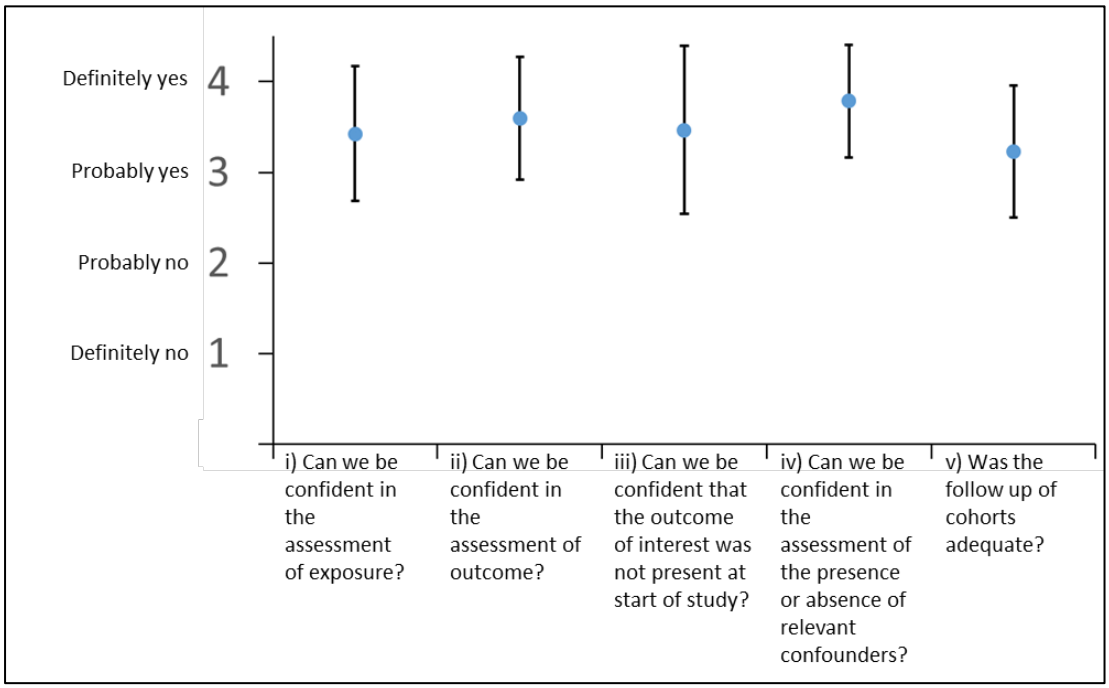


Figure 3. Summary of quality and risk of bias assessment across studies. Values are mean ± SD.

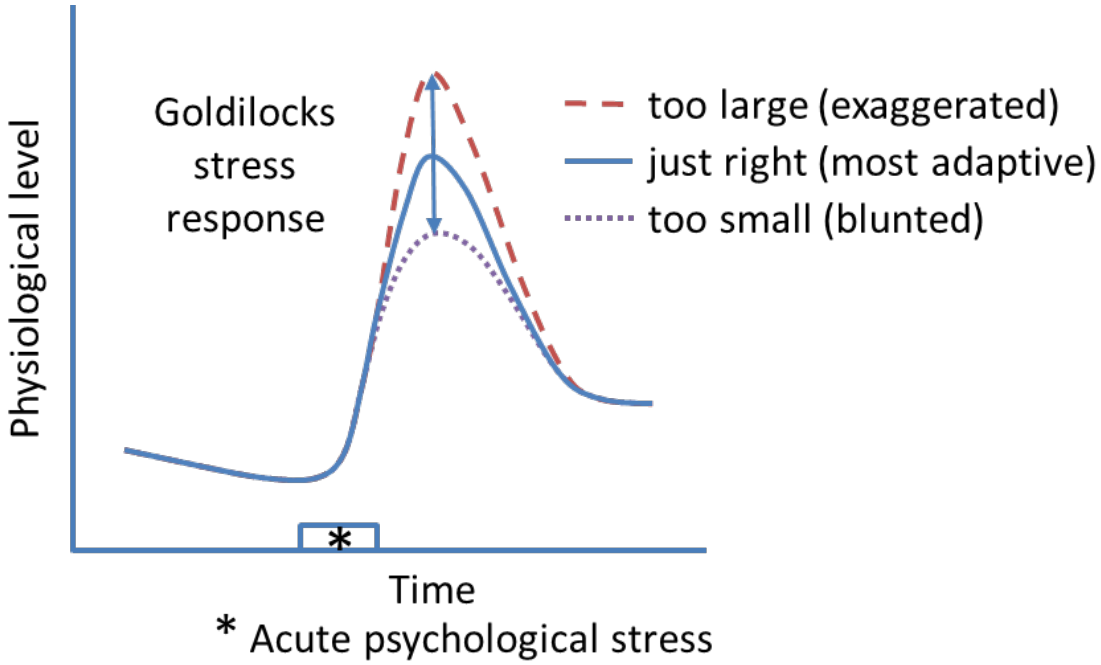


Figure 4. A schematic representation of the Goldilocks stress response. The Goldilocks stress response, represented by the solid/blue line, is an intermediate stress response. It is not too large (exaggerated) and not too small (blunted). Evidence presented in this review shows that it is the most adaptive type of stress response with the lowest risk of future adverse health and disease outcomes.

Table 1. Characteristics of studies included in systematic review (n = 47). Study numbers are defined in Table 1 for use in in Table 2.

No.	Authors, year, country in which study took place	Participants (N, sex, age at baseline[a]), name of cohort study, if relevant	Type of stressor imposed at baseline[b]	Reactivity variable(s) measured at baseline[c]	Follow-up length	Health and/or disease variables measured at follow-up	Significant findings (number of significant findings reported)	Null findings (number of null findings reported)	Study quality, risk of bias assessment, i, ii, iii, iv, v, vi[d]
1	Lawler and Schmied (1992) USA	32 F, 43.8 (27 - 64) y	Structured interview	SBP, DBP, HR	5 y	Frequency and severity of illness in past 12 months	Blunted SBP reactivity at baseline predicted greater illness frequency at follow-up (1)	SBP reactivity did not predict illness severity; DBP and HR reactivity did not predict illness outcomes (5)	4, 4, 4, 4, 2, n
2	Light et al. (1992) USA	30 M, 19.9 ± 1.3 y	Reaction time task with threat of shock (no shock was delivered)	SBP, DBP, HR	10-15 y	Resting SBP, DBP	SBP reactivity at baseline predicted resting SBP at follow-up; DBP and HR reactivity at baseline predicted resting DBP at follow-up (3)	DBP, HR reactivity at baseline did not predict resting SBP at follow-up (2)	4, 4, 4, 4, 2, n
3	Matthews et al. (1993) USA	106 F, 41.8 ± 5.3 y 76 M, 44.1 ± 6.5 y	Mental arithmetic and mirror tracing tasks	SBP, DBP, HR	6.5 ± 0.5 y	Resting SBP, DBP, HR	DBP reactivity to mental arithmetic and mirror tracing at baseline predicted resting DBP at follow-up, as did SBP reactivity to mental arithmetic (3)	SBP reactivity to mirror tracing did not predict resting DBP at follow-up (P<0.10); SBP reactivity to mental arithmetic did not predict resting SBP at follow-up (P<0.07); No results are reported for HR reactivity at baseline or resting HR at follow-up (2)	4, 4, 4, 4, 3, n
4	Brody and Rau (1994)	34 F, 46 M, 31.4 ± 6.0 y	Mental arithmetic[e]	SBP, DBP, HR	1.6 y	Resting SBP, DBP	Blunted HR reactivity at baseline predicted	SBP, DBP, HR reactivity at baseline	2, 4, 4, 2, 3, n

	Germany						greater increase in resting DBP from baseline to follow-up (1)	did not predict change in resting SBP from baseline to follow-up (3)	
5	Carroll et al. (1995) England	1003 M, 44.1 ± 5.9 y, Whitehall II	Mental stress task (Raven's matrices)	SBP, DBP	4.9 ± 0.5 y	Resting SBP, DBP	SBP reactivity at baseline predicted resting SBP at follow-up; DBP reactivity at baseline predicted resting DBP at follow-up (2)	(0)	4, 3, 4, 3, 3, n
6	Brody et al. (1996)[f] Germany	28 F, 27.0 ± 6.5 y 47 M, 27.9 ± 6.2 y	Mental arithmetic	MAP, HR	4 y	Resting MAP	(0)	MAP and HR reactivity at baseline did not predict change in resting MAP from baseline to follow-up (2)	3, 3, 4, 4, 3, n
7	Everson et al. (1996) Finland	508 M, 51.0 ± 6.7 y, Kuopio Ischemic Heart Disease Study	Anticipation of an exercise	SBP, DBP	4.1 y (2.3 - 5.2 y)	Hypertensive status	SBP and DBP reactivity at baseline predicted presence of hypertension at follow-up (2)	(0)	4, 3, 4, 4, 3, n
8	Markovitz et al. (1998) USA	1776 F, 1588 M, 27.0 (20 - 32) y, CARDIA Study	Mirror tracing and video game	SBP, DBP	5 y	Increase in resting SBP, DBP ≥ 8 mmHg, hypertensive status	SBP reactivity to video game at baseline predicted increase in resting SBP (≥ 8 mmHg) in men only; In black men only, new hypertensives had a greater DBP reactivity to video game at baseline than normotensives (2)	SBP reactivity to video game at baseline did not predict increase in resting SBP (≥ 8 mmHg) in women; DBP reactivity to video game did not predict increase in resting DBP (≥ 8 mmHg); SBP, DBP reactivity to mirror tracing did not predict increase in resting SBP, DBP (≥ 8 mmHg);	3, 4, 4, 4, 3, n

								SBP reactivity to video game or mirror tracing did not differ between new hypertensives and normotensives for any race/gender group; In women and white men, DBP reactivity to video game did not differ between new hypertensives and normotensives; DBP reactivity to mirror tracing did not differ between new hypertensives and normotensives for any race/gender group (19)	
9	Carroll et al. (2001)[g] England	796 M, 44.1 ± 5.9 y, Whitehall II	Mental stress task (Raven's matrices)	SBP, DBP	10.8 ± 0.5 y	Resting SBP, DBP, hypertensive status	SBP reactivity at baseline predicted resting SBP at follow-up (1)	SBP, DBP reactivity at baseline did not predict resting DBP at follow-up; SBP, DBP reactivity at baseline did not predict hypertensive status at follow-up (after adjustment for covariates; p=0.07 for DBP reactivity) (4)	4, 3, 4, 3, 3, n
10	Stewart and France (2001) USA	43 F, 30 M, 18 - 20 y	Mental arithmetic	SBP, DBP, HR	3 y	Resting SBP, DBP	HR reactivity at baseline predicted resting SBP at follow-up (1)	SBP, DBP reactivity at baseline did not predict resting SBP at follow-up; SBP, DBP, HR reactivity at baseline did not predict resting DBP at follow-up	3, 4, 3, 4, 2, n

								(5)	
11	Carroll et al. (2003) Scotland	541 F, 449 M, 41.7 ± 14.8 y, West of Scotland Twenty-07 Study	PASAT	SBP, DBP	5.1 ± 0.5 y	Resting SBP, DBP	SBP reactivity at baseline predicted resting SBP at follow-up; SBP, DBP reactivity at baseline predicted 5-y upward drift in resting SBP, DBP, respectively (3)	DBP reactivity at baseline was not associated with resting DBP at follow-up (1)	4, 3, 4, 4, 3, n
12	Fauvel et al. (2003) France	292 F&M for progression to hypertension, 286 F&M for end of follow-up (1-5-y) resting SBP, DBP, 131 F&M for 5-y ambulatory SBP, DBP, 38 (18 - 55) y	Stroop task	SBP, DBP, HR	Annual check-up for up to 5.0 (1-6) y	Progression to hypertension assessed at annual check-up, end of follow-up (1-5-y) resting SBP, DBP, 5-y ambulatory SBP, DBP	(0)	Progression to hypertension did not differ between high-SBP-reactivity and non-high-SBP-reactivity groups; End of follow-up (1-5-y) resting SBP, DBP did not differ between high-SBP-reactivity and non-high-SBP-reactivity groups; 5-y ambulatory SBP, DBP did not differ between high-SBP-reactivity and non-high-SBP-reactivity groups; No results were reported for DBP or HR reactivity (5)	3, 4, 4, 4, 3, n
13	Jennings et al. (2004) Finland	756 M, 42 - 60 y, Kuopio Ischemic Heart Disease Study	Memory, reaction time, reversed – motion tracing and Stroop tasks	SBP, DBP, HR	7 y	Carotid intima-media thickness (IMT)	SBP reactivity at baseline predicted progression of mean carotid IMT, maximal carotid IMT and plaque height from baseline to follow-up (3)	DBP, HR reactivity at baseline did not predict IMT measures at follow-up (6)	4, 4, 4, 4, 4, n

14	Matthews et al. (2004) USA	~4100 F&M, 27 (18 - 30) y, CARDIA Study	Mirror tracing and video game	SBP, DBP	3, 5, 8 and 13 y	Time to occurrence of hypertension	SBP reactivity to mirror tracing and video game at baseline predicted earlier onset of hypertension; DBP reactivity to mirror tracing at baseline predicted earlier onset of hypertension; DBP reactivity to video game at baseline predicted earlier onset of hypertension in men (4)	DBP reactivity to video game at baseline did not predict earlier onset of hypertension in women (1)	4, 4, 4, 4, 4, n
15	Brydon and Steptoe (2005) England	68 F, 51.9 ± 2.9 y, 85 M, 52.6 ± 2.5 y, Whitehall II	Stroop and mirror tracing tasks (average task reactivity was analysed)	SBP, DBP	3.1 ± 0.3 y	Ambulatory SBP, DBP	SBP reactivity at baseline predicted ambulatory SBP at follow-up (1)	DBP reactivity at baseline did not predict ambulatory DBP at follow-up (1)	4, 4, 4, 4, 4, n
16	Steptoe and Marmot (2005) England	98 F, 51.8 ± 2.8 y, 111 M, 52.6 ± 2.6 y, Whitehall II	Stroop and mirror tracing tasks (average task reactivity was analysed)	SBP, DBP, HR	3.1 ± 0.3 y	Resting SBP, DBP	(0)	SBP, DBP, HR reactivity at baseline did not predict resting SBP or DBP at follow-up after adjustment for post-stress recovery values (measured 40–45 min after stress) (6)	4, 2, 4, 4, 4, n
17	Steptoe and Wardle (2005) England	96 F, 51.7 ± 2.7 y, 110 M, 52.7 ± 2.7 y, Whitehall II	Stroop and mirror tracing tasks (average task reactivity was analysed)	SBP, DBP, HR	3.0 y	BMI, WHR change across 3 y	(0)	There was no significant change in BMI across the 3 y (SBP, DBP, HR reactivity at baseline did not predict BMI change);	4, 4, 4, 4, 4, p

								No results are reported for SBP, DBP, HR reactivity at baseline and WHR change (only post-stress recovery results are reported) (3)	
18	Deter et al. (2006) Germany	31 M, 26.6 ± 3.3 y	Manometer test (information-processing task under time pressure)	SBP, DBP, HR	4.8 y	Ambulatory SBP, DBP	SBP, DBP reactivity at baseline predicted ambulatory DBP at follow-up (2)	SBP, DBP reactivity at baseline did not predict ambulatory SBP at follow-up; No results are reported for HR reactivity (2)	4, 4, 4, 4, 4, n
19	Matthews et al. (2006) USA	2816 F&M, 20 - 35 y, CARDIA Study	Mirror tracing and video game	SBP, DBP, HR	13 y	Coronary artery calcification (CAC)	SBP reactivity to video game at baseline predicted CAC at follow-up; Blunted HR reactivity to video game at baseline predicted CAC at follow-up in blacks (2)	DBP reactivity to video game at baseline did not predict CAC at follow-up; Blunted HR reactivity to video game at baseline did not predict CAC at follow-up in whites; SBP, DBP, HR reactivity to mirror tracing at baseline did not predict CAC at follow-up (5)	4, 4, 3, 4, 3, n
20	Stewart et al. (2006) USA	110 F, 106 M, 60.1 ± 4.6 y, Pittsburgh Healthy Heart Project	Marksmanship, visual short-term memory, psychomotor, Stroop, and speech tasks (aggregate task	SBP, DBP, HR	3.2 (2.3-4.3) y	Resting SBP, DBP	(0)	Resting SBP, DBP decreased significantly from baseline to follow-up (SBP, DBP, HR reactivity at baseline did not predict change in resting SBP, DBP	3, 3, 4, 4, 3, n

			reactivity was analysed)					from baseline to follow-up) (6)	
21	Heponiemi et al. (2007) Finland	33 F, 33 M, 28.5 ± 4.7 y[h], Cardiovascular Risk in Young Finns (CRYF) study	Mental arithmetic and speech tasks (average task reactivity was analysed)	HR	2.0 – 2.5 y	Carotid IMT	Blunted HR reactivity at baseline predicted greater carotid IMT at follow-up (1)	(0)	4, 4, 2, 3, 3, n
22	Carroll et al. (2008) Scotland	1272 F&M, 41.8 ± 15.4 y, West of Scotland Twenty-07 Study	PASAT	SBP, DBP, HR	5.5 ± 1.0 y	BMI, WHR, obesity (BMI ≥ 30kg/m^2)	Blunted HR reactivity at baseline predicted increased likelihood of being obese at follow-up (1)	SBP, DBP, HR reactivity at baseline did not predict increase in BMI or WHR from baseline to follow-up; No results are reported for SBP, DBP reactivity at baseline and likelihood of being obese at follow-up (6)	3, 4, 4, 4, 3, n
23	Flaa et al. (2008a) Norway	80 M, 19.3 ± 0.4 y, Military draft screening	Mental arithmetic	Arterial A, NA	18.0 ± 0.9 y	Fasting glucose, HOMA-IR	(0)	A, NA reactivity at baseline did not predict fasting glucose or HOMA-IR at follow-up (4)	4, 4, 3, 4, 4, n
24	Flaa et al. (2008b) Norway	80 M, 19.3 ± 0.4 y, Military draft screening	Mental arithmetic	SBP, DBP, HR, arterial A, NA	18.0 ± 0.9 y	Resting SBP, DBP	When resting SBP was adjusted for, absolute levels of SBP, NA during mental arithmetic positively predicted and absolute levels of A negatively predicted resting SBP at follow-up; When resting DBP was adjusted for, absolute	SBP, DBP, HR, A, NA reactivity at baseline did not predict resting SBP, DBP at follow-up (10)	3, 4, 2, 4, 4, n

							levels of DBP during mental arithmetic positively predicted and absolute levels of A negatively predicted resting DBP at follow-up (5)		
25	Flaa et al. (2008c) Norway	80 M, 19.3 ± 0.4 y, Military draft screening	Mental arithmetic	Arterial A, NA	18.0 ± 0.9 y	BMI, waist circumference, triceps skinfold thickness	Blunted A reactivity at baseline predicted BMI, waist circumference and triceps skinfold thickness at follow-up; NA reactivity at baseline predicted waist circumference at follow-up (4)	NA reactivity at baseline was a near significant predictor of BMI (P=0.08) and triceps skinfold thickness (P=0.08) at follow-up (2)	4, 3, 4, 4, 4, n
26	Phillips et al. (2009) Scotland	1318 F&M, 41.8 ± 15.4 y, West of Scotland Twenty-07 Study	PASAT	SBP, DBP, HR	5.5 ± 1.0 y	Self-reported health	Blunted DBP, HR reactivity at baseline predicted poorer self-reported health at follow-up (2)	SBP reactivity at baseline did not predict self-reported health at follow-up (1)	3, 3, 4, 4, 3, n
27	Carroll et al. (2011)[i] Scotland	645 F, 551 M, 40.7 ± 14.8 y, West of Scotland Twenty-07 Study	PASAT	SBP, DBP	12.4 ± 0.4 y	Resting SBP, DBP, upward drift in SBP, DBP, hypertensive status	SBP reactivity at baseline predicted resting SBP, upward drift in SBP and risk of being hypertensive at follow-up (3)	DBP reactivity at baseline did not predict resting DBP (P=0.06), upward drift in DBP (P=0.06) or risk of being hypertensive at follow-up (3)	3, 4, 4, 4, 4, n
28	Ginty et al. (2011) Scotland	618 - 677 F, 529 - 574 M, 42.2 ± 15.4 y, West of Scotland Twenty-07 Study	PASAT	SBP, DBP, HR	5.5 ± 1.0 y, 12.4 ± 0.4 y	Cognitive function (general intelligence, choice reaction time)	Blunted HR reactivity at baseline predicted lower general intelligence scores and slower choice reaction times at 5-	SBP, DBP reactivity at baseline did not predict general intelligence scores at 5- or 12-y follow-up, choice reaction time	3, 3, 3, 4, 3, n

							and 12-y follow-up and greater decline between assessments for both measures; Blunted SBP reactivity predicted slower choice reaction times at 5-y follow-up (7)	at 12-y follow-up or decline of either measure between assessments; DBP reactivity at baseline did not predict choice reaction time at 5-y follow-up (11)	
29	Phillips et al. (2011a) Scotland	459 F, 393 M, 51.9 ± 9.4 y, West of Scotland Twenty-07 Study	PASAT	SBP, DBP, HR	5.4 ± 0.9 y	Self-reported physical disability (including subscales for locomotion, reaching and stretching, dexterity, seeing, hearing and continence)	Blunted HR reactivity at baseline predicted greater deterioration of physical function over 5 y of follow-up (1)	SBP, DBP reactivity at baseline did not predict change in disability score (2)	3, 4, 4, 4, 3, n
30	Phillips et al. (2011b) Scotland	686 F, 559 M, 42.3 ± 15.5 y, West of Scotland Twenty-07 Study	PASAT	SBP, DBP, HR	5.5 ± 1.0 y	Depression, anxiety symptoms	Blunted HR reactivity at baseline predicted more depression symptoms at follow-up (1)	HR reactivity at baseline did not predict anxiety symptoms at follow up (after adjustment for covariates); SBP, DBP reactivity at baseline did not predict depression or anxiety symptoms at follow-up (5)	3, 4, 4, 4, 3, n
31	Hamer and Steptoe (2012) England	222 F, 257 M, 62.7 ± 5.6 y, Whitehall II	Stroop and mirror tracing tasks (presented sequentially, in random order, as one stressor)	Salivary cortisol (0, 10, 30, 55, 85 min; change score and AUCi)	3.0 (1.8 - 3.6) y	Incident hypertension	Salivary cortisol reactivity at baseline (both change score and AUCi) predicted risk of hypertension at follow-up (1)	(0)	3, 4, 4, 4, 4, n

32	Hamer et al. (2012) England	213 F, 253 M, 62.7 ± 5.6 y, Whitehall II	Stroop and mirror tracing tasks (presented sequentially, in random order, as one stressor)	SBP, DBP, salivary cortisol (0, 10, 30, 55, 85 min; change score)	3.0 (1.8 - 3.6) y	Risk of CAC progression, relative CAC change	Salivary cortisol reactivity at baseline predicted risk of CAC progression and relative CAC change (2)	SBP, DBP reactivity at baseline did not predict risk of CAC progression (2)	3, 4, 4, 4, 4, n
33	Phillips et al. (2012) The Netherlands	240 F, 220 M, 58.3 ± 0.9 y, Dutch Famine Birth Cohort	Stroop, mirror tracing and speech tasks (aggregate data were used)	SBP, DBP, HR, salivary cortisol (-15, 0, 11, 22, 37, 47, 57 min; change score)	4 – 7 y	BMI, obesity (BMI ≥ 30kg/m^2)	Blunted HR reactivity at baseline predicted increased risk of becoming or remaining obese between baseline and follow-up (1)	SBP, DBP, cortisol reactivity at baseline did not predict change in BMI; Results for other exposure/outcome combinations are not reported (3)	4, 1, 4, 4, 2, n
34	Shaffer et al. (2012) Canada	719 - 724 F, 743 - 746 M, 45.3 ± 17.9 y, Nova Scotia Health Survey (1995)	Anger provocation interview	SBP, DBP	10 y	Fatal or nonfatal incident CVD events	(0)	SBP, DBP reactivity at baseline did not predict risk of incident CVD events during 10 y of follow-up (2)	1, 3, 3, 4, 3, n
35	Senan and Petrosyan (2014) India	86 F, 260 M, 41 - 54 y	Anticipation of exercise	HR	Up to 7 y	Sudden coronary death, non-sudden coronary death, death from any cause	HR reactivity at baseline predicted risk for sudden coronary death and death from any cause during the 7 y of follow-up (2)	HR reactivity at baseline did not predict risk for non-sudden coronary death during the 7 y of follow-up (1)	1, 2, 3, 1, 1, p
36	Gentile et al. (2015) Canada	81 F, 55 M, 41 ± 11.5 y	A neutral reading task, two role plays, and a non-scripted debate (aggregate data were used)	SBP, DBP, HR	2.9 ± 0.3 y	Metabolic burden (number of metabolic syndrome parameters for which participants were in highest quartile (lowest for HDL)), risk of having metabolic syndrome	(0)	SBP, DBP, HR reactivity at baseline did not predict metabolic burden or risk of having metabolic syndrome at follow-up (6)	3, 4, 4, 4, 3, n

37	Paananen et al. (2015) Australia	198 – 222 F, 198 – 215 M, 18.3 ± 0.3 y, Western Australian Pregnancy Cohort (Raine) Study	TSST	Plasma total cortisol (0, 15, 25, 35, 45, 60, 75, 105 min; cluster analysis)	4 y	Musculoskeletal pain alone (any pain, high pain), musculoskeletal pain combined with increased sensitivity to cold, pressure pain	Blunted plasma cortisol reactivity at baseline predicted greater likelihood of musculoskeletal pain (any pain, high pain) in females with greater sensitivity to cold pain; Blunted plasma cortisol reactivity at baseline predicted greater likelihood of high musculoskeletal pain in females with greater sensitivity to pressure pain (3)	Plasma cortisol reactivity at baseline did not predict musculoskeletal pain alone in females or any musculoskeletal pain in females with greater sensitivity to pressure pain; Plasma cortisol reactivity at baseline did not predict any musculoskeletal pain outcomes in males (9)	3, 4, 1, 4, 3, n
38	Steudte-Schmiedgen et al. (2015) Germany	80 M, 27.8 ± 5.9 y, Prevalence, Incidence and Determinants of PTSD and Other Mental Disorders (PID-PTSD+3) Study, German federal defence force	TSST	Salivary cortisol (-1, 16, 25, 35 min; AUCi)	1.6 y, baseline was 1 – 3 months before deployment, deployment was 5.2 ± 1.5 months, and on average, follow-up was 12 months after return	PTSD symptomatology after expose to new-onset traumatic events during military deployment	Blunted salivary cortisol reactivity at baseline predicted a greater increase in PTSD symptomatology in soldiers who had experienced new-onset traumatic events (1)	(0)	3, 4, 4, 2, 2, n
39	Yuenyongchaiwat et al. (2015) Thailand	75 F, 26 M, 31.3 ± 9.5 y	Mental arithmetic and speech tasks	SBP, DBP, HR	1.1 (1.0 – 1.3) y	Resting SBP, DBP	SBP reactivity to mental arithmetic at baseline predicted resting SBP at follow up (1)	SBP, DBP, HR reactivity at baseline did not predict resting DBP at follow-up; Results for other exposure/outcome combinations are not reported (3)	4, 4, 3, 4, 4, n

40	Brindle et al. (2016) The Netherlands	230 F, 208 M, 58.4 (55 - 60) y, Dutch Famine Birth Cohort	Stroop, mirror tracing and speech tasks (aggregate data were used)	SBP, DBP, HR (cluster analysis was used)	5.5 ± 0.6 y	Hypertensive status	Membership of a cluster with high SBP, DBP reactivity combined with moderate HR reactivity at baseline predicted greater risk of hypertension at follow-up (1)	Membership of clusters with high reactivity of SBP, DBP & HR or low reactivity of SBP, DBP & HR at baseline did not predict risk of hypertension at follow-up (2)	4, 3, 3, 4, 3, n
41	Ronaldson et al. (2016) England	84 F, 37 M, 63.5 ± 5.7 y Whitehall II	Stroop and mirror tracing tasks (presented sequentially, in random order, as one stressor)	Salivary cortisol (0, 10, 30, 55, 85 min; change score)	3 y	Regulatory T cell percentage, responder T cell percentage, physical health, mental health, depression symptoms	Blunted salivary cortisol reactivity at baseline predicted greater regulatory T cell and lower responder T cell percentage at follow-up; In turn, higher regulatory T cell percentage was associated with worse physical and mental health and greater depressive symptomatology (2)	(0)	3, 4, 1, 4, 3, n
42	Steptoe et al. (2016) England	299 F, 307 M, 59.1 ± 6.7 y, Whitehall II	Stroop and mirror tracing tasks (average task reactivity was analysed)	SBP, DBP	7.9 ± 3.4 y	Incident hypertension	SBP, DBP reactivity at baseline predicted risk of hypertension at follow-up (2)	(0)	4, 3, 4, 4, 4, n
43	Yuichiro et al. (2016) USA	1698 F, 1323 M 27.1 ± 3.6 y, CARDIA Study	Mirror tracing and video game	SBP, DBP	23 y	Cognitive function (psychomotor speed, verbal memory and executive function)	Blunted SBP reactivity to mirror tracing and video game at baseline predicted worse psychomotor speed and executive function at follow-up;	SBP reactivity at baseline did not predict verbal memory at follow-up; DBP reactivity to mirror tracing at baseline did not predict verbal	4, 4, 1, 4, 4, n

							Blunted DBP reactivity to mirror tracing at baseline predicted worse psychomotor speed at follow-up; Blunted DBP reactivity to video game at baseline predicted worse verbal memory at follow-up (6)	memory or executive function at follow-up; DBP reactivity to the video game did not predict psychomotor speed or executive function at follow-up (6)	
44	Zhu et al. (2016) Australia	390 F, 411 M for salivary cortisol, 291 F, 357 M for plasma cortisol, 18.3 ± 0.3 y, Western Australian Pregnancy Cohort (Raine) Study	TSST	Salivary cortisol (0, 15, 35, 105 min; AUCi) and total plasma cortisol (0, 15, 25, 35, 45, 60, 75, 105 min; AUCi)	2 y	Total body bone mineral content (BMC), bone area, bone mineral density (BMD)	Blunted salivary cortisol reactivity at baseline predicted lower BMD, BMC at follow-up in males (2)	Salivary cortisol reactivity at baseline did not predict BMD, BMC at follow-up in females or bone area at follow-up in either sex; Total plasma cortisol reactivity at baseline did not predict any bone outcomes at follow-up (10)	3, 4, 1, 4, 3, n
45	Steptoe et al. (2017) England	215 F, 196 M, 54 - 76 y Whitehall II	Stroop and mirror tracing tasks (presented sequentially, in random order, as one stressor)	Salivary cortisol (0, 10, 30, 55, 85 min; change score)	3 y	Leukocyte telomere length	Salivary cortisol reactivity at baseline predicted shorter telomere length at follow-up (1)	(0)	4, 4, 4, 4, 4, n
46	Yuenyongchaiwat (2017)[j] Thailand	74 F, 21 M, 32.2 ± 10.2 y	Mental arithmetic and speech tasks	SBP, DBP, HR	3.4 ± 0.1 y	Resting SBP, DBP	SBP reactivity to mental arithmetic at baseline predicted resting SBP at follow-up (1)	SBP reactivity to speech task at baseline did not predict resting SBP at follow-up; DBP reactivity at baseline did not predict resting DBP at follow-up;	4, 4, 3, 4, 4, n

								Results for other exposure/outcome combinations are not reported (3)	
47	Yuenyongchaiwat and Sheffield (2017) Thailand	75 F, 27 M, 31.9 ± 10.0 y[h]	Mental arithmetic and speech tasks	SBP, DBP, HR	3.4 (3.2 - 3.6) y	Depression, anxiety symptoms	Blunted SBP reactivity to mental arithmetic at baseline predicted anxiety symptoms at follow-up (1)	SBP reactivity to speech task did not predict anxiety at follow-up; DBP, HR reactivity at baseline did not predict anxiety at follow-up; SBP, DBP, HR reactivity at baseline did not predict depression symptoms at follow-up (11)	4, 4, 4, 4, 4, n

Data are presented as mean ± SD or mean (range) or mean or range; F = female; M = male; y = years; SBP = systolic blood pressure; DBP = diastolic blood pressure; MAP = mean arterial pressure; HR = heart rate; IMT = intima-media thickness; BMI = body mass index; CAC = coronary artery calcification; PASAT = paced auditory serial addition test; WHR = waist-hip ratio; A = adrenaline; NA = noradrenaline; HOMA-IR = homeostasis model assessment of insulin resistance; CVD = cardiovascular disease; TSST = Trier Social Stress Test; Stroop task = Stroop colour–word interference task; Mirror tracing = Mirror image star-tracing task; PTSD = Post Traumatic Stress Disorder

[a] N is reported as the number of participants who were included in the final longitudinal analysis, not the number who entered the study or the number who may have been included in concurrent cross-sectional, or other, analyses, if applicable. Age is reported as mean, range or mean±SD, where available.

[b] Only stressors included in the final analysis are included here. Where studies imposed both physical and psychological stressors, only the psychological stressors are included in this systematic review.

[c] Includes sample timing and method of reactivity calculation for cortisol. Time = 0 minutes indicates the commencement of the stressor(s).

[d] Study quality and risk of bias assessment utilised the following questions and answer codes:
i) Can we be confident in the assessment of exposure?
ii) Can we be confident in the assessment of outcome?
iii) Can we be confident that the outcome of interest was not present at start of study?
iv) Can we be confident in the assessment of the presence or absence of relevant confounders?
v) Was the follow up of cohorts adequate?
vi) Was there evidence of any other sources of bias in reporting?

4 = Definitely yes (low risk of bias); 3 = Probably yes; 2 = Probably no; 1 = Definitely no (high risk of bias); n = no; p = possibly

[e] This paper also did measurements of pain at baseline but does not indicate when this occurred in relation to when stress testing occurred, therefore, it is not clear if pain testing interfered with stress reactivity testing.

[f] This was a 4-y follow-up of the same cohort that was earlier followed-up at 1.6 y (Brody and Rau, 1994)

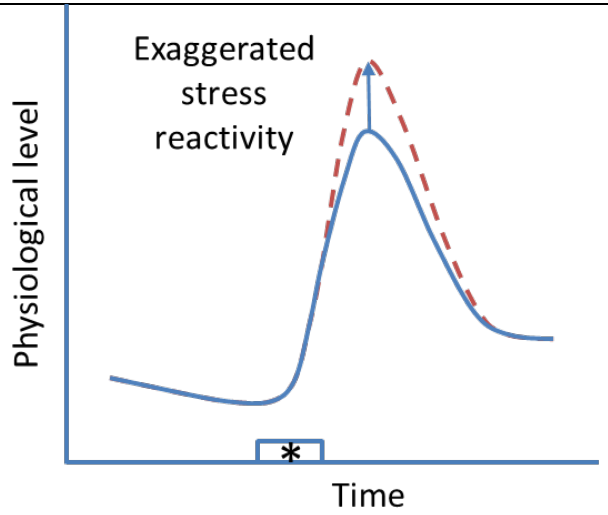
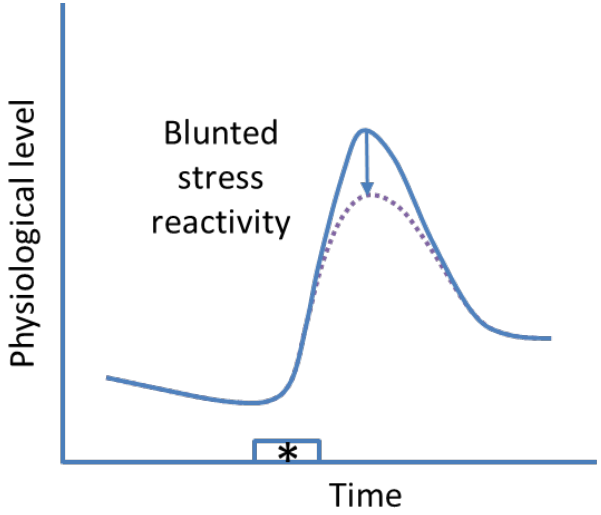
[g] This was a 10.8-y follow-up of the same cohort that was earlier followed-up at 4.9 y (Carroll et al., 1995)

[h] Age at baseline of study sample calculated as age at follow-up minus follow-up length.

[i] This was a 12-y follow-up of the same cohort that was earlier followed-up at 5 y (Carroll et al., 2003).

[j] This was a 3.4-y follow-up of the same cohort that was earlier followed-up at 1.1 y (Yuenyongchaiwat et al., 2015).

Table 2. Stress reactivity at baseline[a] and its relationship to health and disease outcomes at follow-up: Synthesis of main significant findings[b]

	SAM system reactivity	HPA axis reactivity
	**Exaggerated SAM system reactivity** • (SBP, DBP, HR) Greater resting SBP, DBP [2; 3; 5; 8; 9; 10; 11; 24; 27; 39; 46] • (SBP, DBP) Greater ambulatory SBP, DBP [15; 18] • (SBP, DBP) Greater risk of hypertension [7; 27; 40; 42] • (SBP, DBP) Earlier onset of hypertension [14] • (SBP) Greater progression of IMT [13] • (SBP) Greater odds of having CAC [19] • (NA) Greater waist circumference [25] • (HR) Increased risk for sudden coronary death and death from any cause [35]	**Exaggerated HPA axis reactivity** • (SC) Greater risk of hypertension [31] • (SC) Greater CAC progression [32] • (SC) More rapid telomere attrition [45]
	**Blunted SAM system reactivity** • (SBP) Greater illness frequency [1] • (HR) Greater increases in resting DBP [4] • (HR) Greater odds of having CAC in blacks (but not whites) [19] • (HR) Greater carotid IMT [21] • (A) Greater BMI, waist circumference and triceps skinfold thickness [25] • (HR) Increased likelihood of being obese [22; 33] • (DBP, HR) Poorer self-reported health [26] • (HR) Greater progression of self-reported physical disability [29] • (HR) More depression symptoms [30] • (HR, SBP, DBP) Reduced cognitive ability [28; 43] • (SBP) More anxiety symptoms [47]	**Blunted HPA axis reactivity** • (PC) Greater musculoskeletal pain in pain-sensitive individuals in females (but not males) [37] • (SC) More PTSD symptomology after exposure to new-onset traumatic events [38] • (SC) Greater regulatory T-Cell percentage (which in turn, was associated with poorer physical and mental health and more depressive symptomatology) [41] • (SC) Lower bone mineral content and bone mineral density in males (but not females) [44]

SAM system = sympatho-adrenal medullary system; HPA axis = hypothalamo-pituitary adrenal axis; SBP = systolic blood pressure; DBP = diastolic blood pressure; HR = heart rate; NA = noradrenaline; A = adrenaline; SC = salivary cortisol; PC = plasma cortisol; IMT = intima-media thickness; CAC = coronary artery calcification; BMI = body mass index; * acute psychological stress

[a] Stress reactivity variables measured at baseline are indicated in parentheses ahead of each outcome measured at follow-up.

[b] Superscript numbers refer to study numbers defined in Table 1.

References

Ali, N., Pruessner, J.C., 2012. The salivary alpha amylase over cortisol ratio as a marker to assess dysregulations of the stress systems. Physiology & Behavior 106, 65-72.
Brindle, R.C., Ginty, A.T., Jones, A., Phillips, A.C., Roseboom, T.J., Carroll, D., Painter, R.C., De Rooij, S.R., 2016. Cardiovascular reactivity patterns and pathways to hypertension: A multivariate cluster analysis. J. Hum. Hypertens. 30, 755-760.
Brody, S., Rau, H., 1994. Behavioral and psychophysiological predictors of self-monitored 19 month blood pressure change in normotensives. J. Psychosomat. Res. 38, 885-891.
Brody, S., Veit, R., Rau, H., 1996. Neuroticism but not cardiovascular stress reactivity is associated with less longitudinal blood pressure increase. Pers. Individ. Differ. 20, 375-380.
Brunton, L.L.e., Knollmann, B.C.e., Hilal-Dandan, R.e., 2018. Goodman & Gilman's : The Pharmacological Basis of Therapeutics, 13e. McGraw-Hill Education LLC, New York, N.Y.
Brydon, L., Steptoe, A., 2005. Stress-induced increases in interleukin-6 and fibrinogen predict ambulatory blood pressure at 3-year follow-up. Journal of Hypertension 23, 1001-1007.
Bunea, I.M., Szentagotai-Tatar, A., Miu, A.C., 2017. Early-life adversity and cortisol response to social stress: a meta-analysis. Transl. Psychiatr. 7, 1274.
Carroll, D., Davey Smith, G., Shipley, M.J., Steptoe, A., Brunner, E.J., Marmot, M.G., 2001. Blood pressure reactions to acute psychological stress and future blood pressure status: A 10-year follow-up of men in the Whitehall II study. Psychosomatic Medicine 63, 737-743.
Carroll, D., Ginty, A.T., Whittaker, A.C., Lovallo, W.R., de Rooij, S.R., 2017. The behavioural, cognitive, and neural corollaries of blunted cardiovascular and cortisol reactions to acute psychological stress. Neurosci. Biobehav. Rev. 77, 74-86.
Carroll, D., Phillips, A.C., Der, G., 2008. Body mass index, abdominal adiposity, obesity, and cardiovascular reactions to psychological stress in a large community sample. Psychosomatic Medicine 70, 653-660.
Carroll, D., Phillips, A.C., Der, G., Hunt, K., Benzeval, M., 2011. Blood pressure reactions to acute mental stress and future blood pressure status: Data from the 12-year follow-up of the West of Scotland Study. Psychosomatic Medicine 73, 737-742.
Carroll, D., Ring, C., Hunt, K., Ford, G., MacIntyre, S., 2003. Blood Pressure Reactions to Stress and the Prediction of Future Blood Pressure: Effects of Sex, Age, and Socioeconomic Position. Psychosomatic Medicine 65, 1058-1064.
Carroll, D., Smith, G.D., Sheffield, D., Shipley, M.J., Marmot, M.G., 1995. Pressor reactions to psychological stress and prediction of future blood pressure: Data from the Whitehall II study. Br. Med. J. 310, 771-776.
Chida, Y., Steptoe, A., 2010. Greater Cardiovascular Responses to Laboratory Mental Stress Are Associated With Poor Subsequent Cardiovascular Risk Status A Meta-Analysis of Prospective Evidence. Hypertension 55, 1026-U1368.
Chrousos, G.P., 2009. Stress and disorders of the stress system. Nat. Rev. Endocrinol. 5, 374-381.
Cinciripini, P.M., 1986. Cognitive stress and cardiovascular reactivity. 1. Relationship to hypertension. Am. Heart J. 112, 1044-1050.
Deter, H.-C., Micus, C., Wagner, M., Sharma, A.M., Buchholz, K., 2006. Salt sensitivity, anxiety, and irritability predict blood pressure increase over five years in healthy males. Clin. Exp. Hypertens. 28, 17-27.
Dhillo, W.S., Kong, W.M., Le Roux, C.W., Alaghband-Zadeh, J., Jones, J., Carter, G., Mendoza, N., Meeran, K., O'Shea, D., 2002. Cortisol-binding globulin is important in the interpretation of dynamic tests of the hypothalamic-pituitary-adrenal axis. European Journal of Endocrinology 146, 231-235.
Everson, S.A., Kaplan, G.A., Goldberg, D.E., Salonen, J.T., 1996. Anticipatory blood pressure response to exercise predicts future high blood pressure in middle-aged men. Hypertension 27, 1059-1064.

Fauvel, J.P., M'Pio, I., Quelin, P., Rigaud, J.P., Laville, M., Ducher, M., 2003. Neither Perceived Job Stress Nor Individual Cardiovascular Reactivity Predict High Blood Pressure. Hypertension 42, 1112-1116.
Flaa, A., Aksnes, T.A., Kjeldsen, S.E., Eide, I., Rostrup, M., 2008a. Increased sympathetic reactivity may predict insulin resistance: an 18-year follow-up study. Metabolism. 57, 1422-1427.
Flaa, A., Eide, I.K., Kjeldsen, S.E., Rostrup, M., 2008b. Sympathoadrenal stress reactivity is a predictor of future blood pressure: An 18-year follow-up study. Hypertension 52, 336-341.
Flaa, A., Sandvik, L., Kjeldsen, S.E., Eide, I.K., Rostrup, M., 2008c. Does sympathoadrenal activity predict changes in body fat? An 18-y follow-up study. Am. J. Clin. Nutr. 87, 1596-1601.
Gasperin, D., Netuveli, G., Dias-Da-Costa, J.S., Pattussi, M.P., 2009. Effect of psychological stress on blood pressure increase: a meta-analysis of cohort studies. Cad. Saude Publica 25, 715-726.
Gentile, C., Dragomir, A.I., Solomon, C., Nigam, A., D'Antono, B., 2015. Sex differences in the prediction of metabolic burden from physiological responses to stress. Annals Of Behavioral Medicine: A Publication Of The Society Of Behavioral Medicine 49, 112-127.
Ginty, A.T., Phillips, A.C., Der, G., Deary, I.J., Carroll, D., 2011. Heart rate reactivity is associated with future cognitive ability and cognitive change in a large community sample. Int. J. Psychophysiol. 82, 167-174.
Gozansky, W.S., Lynn, J.S., Laudenslager, M.L., Kohrt, W.M., 2005. Salivary cortisol determined by enzyme immunoassay is preferable to serum total cortisol for assessment of dynamic hypothalamic-pituitary-adrenal axis activity. Clinical Endocrinology 63, 336-341.
Greenhalgh, T., 1997. How to read a paper: Getting your bearings (deciding what the paper is about). Bmj-British Medical Journal 315, 243-246.
Hamer, M., Endrighi, R., Venuraju, S.M., Lahiri, A., Steptoe, A., 2012. Cortisol responses to mental stress and the progression of coronary artery calcification in healthy men and women. PLoS One 7, e31356-e31356.
Hamer, M., Steptoe, A., 2012. Cortisol responses to mental stress and incident hypertension in healthy men and women. Journal of Clinical Endocrinology and Metabolism 97, E29-E34.
Henley, D.E., Leendertz, J.A., Russell, G.M., Wood, S.A., Taheri, S., Woltersdorf, W.W., Lightman, S.L., 2009. Development of an automated blood sampling system for use in humans. J. Med. Eng. Technol. 33, 199-208.
Heponiemi, T., Elovainio, M., Pulkki, L., Puttonen, S., Raitakari, O., Keltikangas-Järvinen, L., 2007. Cardiac autonomic reactivity and recovery in predicting carotid atherosclerosis: the Cardiovascular Risk in Young Finns study. Health Psychol. 26, 13-21.
Jennings, J.R., Kamarck, T.W., Everson-Rose, S.A., Kaplan, G.A., Manuck, S.B., Salonen, J.T., 2004. Exaggerated blood pressure responses during mental stress are prospectively related to enhanced carotid atherosclerosis in middle-aged Finnish men. Circulation 110, 2198-2203.
Jones, A., McMillan, M.R., Jones, R.W., Kowalik, G.T., Steeden, J.A., Deanfield, J.E., Pruessner, J.C., Taylor, A.M., Muthurangu, V., 2012. Adiposity Is Associated with Blunted Cardiovascular, Neuroendocrine and Cognitive Responses to Acute Mental Stress. PLoS One 7, 7.
Kirschbaum, C., Hellhammer, D.H., 1994. Salivary cortisol in psychoneuroendocrine research: recent developments and applications. Psychoneuroendocrinology 19, 313-333.
Krantz, D.S., Manuck, S.B., 1984. Acute psychophysiologic reactivity and risk of cardiovascular-disease - a review and methodologic critique. Psychological Bulletin 96, 435-464.
Lawler, K.A., Schmied, L.A., 1992. A prospective study of women's health: the effects of stress, hardiness, locus of control, Type A behavior, and physiological reactivity. Women Health 19, 27-41.
Light, K.C., 1981. Cardiovascular-responses to effortful active coping - implications for the role of stress in hypertension development. Psychophysiology 18, 216-225.
Light, K.C., 1987. Psychosocial precursors of hypertension - experimental-evidence. Circulation 76, 67-76.

Light, K.C., Dolan, C.A., Davis, M.R., Sherwood, A., 1992. Cardiovascular responses to an active coping challenge as predictors of blood pressure patterns 10 to 15 years later. Psychosomatic Medicine 54, 217-230.
Lightman, S.L., Conway-Campbell, B.L., 2010. The crucial role of pulsatile activity of the HPA axis for continuous dynamic equilibration. Nature Reviews Neuroscience 11, 710-718.
Lindsay, E.K., Young, S.Z., Smyth, J.M., Brown, K.W., Creswell, J.D., 2018. Acceptance lowers stress reactivity: Dismantling mindfulness training in a randomized controlled trial. Psychoneuroendocrinology 87, 63-73.
Lovallo, W.R., Enoch, M.A., Sorocco, K.H., Vincent, A.S., Acheson, A., Cohoon, A.J., Hodgkinson, C.A., Goldman, D., 2017. Joint Impact of Early Life Adversity and COMT Val158Met (rs4680) Genotypes on the Adult Cortisol Response to Psychological Stress. Psychosomatic Medicine 79, 631-637.
Markovitz, J.H., Raczynski, J.M., Wallace, D., Chettur, V., Chesney, M.A., 1998. Cardiovascular reactivity to video game predicts subsequent blood pressure increases in young men: The CARDIA study. Psychosomatic Medicine 60, 186-191.
Matthews, K.A., Katholi, C.R., McCreath, H., Whooley, M.A., Williams, D.R., Zhu, S., Markovitz, J.H., 2004. Blood pressure reactivity to psychological stress predicts hypertension in the CARDIA study. Circulation 110, 74-78.
Matthews, K.A., Woodall, K.L., Allen, M.T., 1993. Cardiovascular reactivity to stress predicts future blood pressure status. Hypertension 22, 479-485.
Matthews, K.A., Zhu, S., Tucker, D.C., Whooley, M.A., 2006. Blood pressure reactivity to psychological stress and coronary calcification in the coronary artery risk development in young adults study. Hypertension 47, 391-395.
Miller, R., Plessow, F., Kirschbaum, C., Stalder, T., 2013. Classification Criteria for Distinguishing Cortisol Responders From Nonresponders to Psychosocial Stress: Evaluation of Salivary Cortisol Pulse Detection in Panel Designs. Psychosomatic Medicine 75, 832-840.
Miller, R., Wojtyniak, J.G., Weckesser, L.J., Alexander, N.C., Engert, V., Lehr, T., 2018. How to disentangle psychobiological stress reactivity and recovery: A for comparison of model-based and non-compartmental analyses of cortisol concentrations. Psychoneuroendocrinology 90, 194-210.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., Group, T.P., 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. PLoS Med 6, e1000097.
Mucke, M., Ludyga, S., Colledge, F., Gerber, M., 2018. Influence of Regular Physical Activity and Fitness on Stress Reactivity as Measured with the Trier Social Stress Test Protocol: A Systematic Review. Sports Med. 48, 2607-2622.
Obrist, P.A., 1976. Cardiovascular-behavioral interaction - As it appears today. Psychophysiology 13, 95-107.
Obrist, P.A., 1981. Cardiovascular psychophysiology : a perspective. Plenum Press, New York.
Ouzzani, M., Hammady, H., Fedorowicz, Z., Elmagarmid, A., 2016. Rayyan-a web and mobile app for systematic reviews. Systematic reviews 5, 210.
Paananen, M., O'Sullivan, P., Straker, L., Beales, D., Coenen, P., Karppinen, J., Pennell, C., Smith, A., 2015. A low cortisol response to stress is associated with musculoskeletal pain combined with increased pain sensitivity in young adults: a longitudinal cohort study. Arthrit. Res. Ther. 17, 355-355.
Phillips, A.C., Der, G., Carroll, D., 2009. Self-reported health and cardiovascular reactions to psychological stress in a large community sample: Cross-sectional and prospective associations. Psychophysiology 46, 1020-1027.
Phillips, A.C., Der, G., Shipton, D., Benzeval, M., 2011a. Prospective associations between cardiovascular reactions to acute psychological stress and change in physical disability in a large community sample. Int. J. Psychophysiol. 81, 332-337.
Phillips, A.C., Hughes, B.M., 2011. Introductory paper: Cardiovascular reactivity at a crossroads: Where are we now? Biological Psychology 86, 95-97.

Phillips, A.C., Hunt, K., Der, G., Carroll, D., 2011b. Blunted cardiac reactions to acute psychological stress predict symptoms of depression five years later: Evidence from a large community study. Psychophysiology 48, 142-148.
Phillips, A.C., Roseboom, T.J., Carroll, D., de Rooij, S.R., 2012. Cardiovascular and cortisol reactions to acute psychological stress and adiposity: Cross-sectional and prospective associations in the Dutch Famine Birth Cohort Study. Psychosomatic Medicine 74, 699-710.
Rohleder, N., 2014. Stimulation of Systemic Low-Grade Inflammation by Psychosocial Stress. Psychosomatic Medicine 76, 181-189.
Ronaldson, A., Gazali, A.M., Zalli, A., Kaiser, F., Thompson, S.J., Henderson, B., Steptoe, A., Carvalho, L., 2016. Increased percentages of regulatory T cells are associated with inflammatory and neuroendocrine responses to acute psychological stress and poorer health status in older men and women. Psychopharmacology 233, 1661-1668.
Senan, M., Petrosyan, A., 2014. Relationship between emotional stress and cardiovascular events. Georgian Medical News, 19-25.
Shaffer, J.A., Wasson, L.T., Davidson, K.W., Schwartz, J.E., Kirkland, S., Shimbo, D., 2012. Blood pressure reactivity to an anger provocation interview does not predict incident cardiovascular disease events: The Canadian Nova Scotia health survey (NSHS95) prospective population study. International Journal of Hypertension 2012, 1-8.
Skoluda, N., La Marca, R., Gollwitzer, M., Muller, A., Limm, H., Marten-Mittag, B., Gundel, H., Angerer, P., Nater, U.M., 2017. Long-term stability of diurnal salivary cortisol and alpha-amylase patterns. Physiology & Behavior 175, 1-8.
Steptoe, A., 2000. Psychosocial factors in the development of hypertension. Annals of Medicine 32, 371-375.
Steptoe, A., Hamer, M., Lin, J., Blackburn, E.H., Erusalimsky, J.D., 2017. The Longitudinal Relationship Between Cortisol Responses to Mental Stress and Leukocyte Telomere Attrition. The Journal Of Clinical Endocrinology And Metabolism 102, 962-969.
Steptoe, A., Kivimäki, M., Lowe, G., Rumley, A., Hamer, M., 2016. Blood pressure and fibrinogen responses to mental stress as predictors of incident hypertension over an 8-year period. nn. Behav. Med. 50, 898-906.
Steptoe, A., Marmot, M., 2005. Impaired cardiovascular recovery following stress predicts 3-year increases in blood pressure. Journal Of Hypertension 23, 529-536.
Steptoe, A., Wardle, J., 2005. Cardiovascular stress responsivity, body mass and abdominal adiposity. International Journal of Obesity 29, 1329-1337.
Steudte-Schmiedgen, S., Stalder, T., Schönfeld, S., Wittchen, H.-U., Trautmann, S., Alexander, N., Miller, R., Kirschbaum, C., 2015. Hair cortisol concentrations and cortisol stress reactivity predict PTSD symptom increase after trauma exposure during military deployment. Psychoneuroendocrinology 59, 123-133.
Stewart, J.C., France, C.R., 2001. Cardiovascular recovery from stress predicts longitudinal changes in blood pressure. Biological Psychology 58, 105-120.
Stewart, J.C., Janicki, D.L., Kamarack, T.W., 2006. Cardiovascular reactivity to and recovery from psychological challenge as predictors of 3-year change in blood pressure. Health Psychol. 25, 111-118.
Strahler, J., Skoluda, N., Kappert, M.B., Nater, U.M., 2017. Simultaneous measurement of salivary cortisol and alpha-amylase: Application and recommendations. Neurosci. Biobehav. Rev. 83, 657-677.
Taylor, T.R., Kamarck, T.W., Dianzumba, S., 2003. Cardiovascular reactivity and left ventricular mass: An integrative review. nn. Behav. Med. 26, 182-193.
Throne, L.C., Bartholomew, J.B., Craig, J., Farrar, R.P., 2000. Stress reactivity in fire fighters: An exercise intervention. Int. J. Stress Manage. 7, 235-246.

Treiber, F.A., Kamarck, T., Schneiderman, N., Sheffield, D., Kapuku, G., Taylor, T., 2003. Cardiovascular reactivity and development of preclinical and clinical disease states. Psychosomatic Medicine 65, 46-62.
West, S.G., Gebauer, S.K., Kay, C.D., Bagshaw, D.M., Savastano, D.M., Diefenbach, C., Kris-Etherton, P.M., 2012. Diets Containing Pistachios Reduce Systolic Blood Pressure and Peripheral Vascular Responses to Stress in Adults With Dyslipidemia. Hypertension 60, 58-63.
Wust, S., van Rossum, E.F.C., Federenko, I.S., Koper, J.W., Kumsta, R., Hellhammer, D.H., 2004. Common polymorphisms in the glucocorticoid receptor gene are associated with adrenocortical responses to psychosocial stress. Journal of Clinical Endocrinology & Metabolism 89, 565-573.
Yuenyongchaiwat, K., 2017. Cardiovascular Response to Mental Stress Tests and the Prediction of Blood Pressure. Indian Journal Of Psychological Medicine 39, 413-417.
Yuenyongchaiwat, K., Sheffield, D., 2017. Blunted cardiovascular reactions are a predictor of negative health outcomes: A prospective cohort study. J. Appl. Biobehav. Res. 22, 1-11.
Yuenyongchaiwat, K., Sheffield, D., Baker, I., Maratos, F., 2015. Hemodynamic responses to active and passive coping tasks and the prediction of future blood pressure in Thai participants: A preliminary prospective cohort study. J. Psych. Res. 57, 288-299.
Yuichiro, Y., Hongyan, N., Reis, J.P., Lewis, C.E., Launer, L.J., Bryan, R.N., Yaffe, K., Sidney, S., Albanese, E., Greenland, P., Lloyd-Jones, D., Kiang, L., 2016. Blood Pressure Reactivity to Psychological Stress in Young Adults and Cognition in Midlife: The Coronary Artery Risk Development in Young Adults (CARDIA) Study. Journal of the American Heart Association 5, 1-9.
Zhu, K., Henley, D., Pennell, C., Herbison, C.E., Mountain, J., Lye, S., Walsh, J.P., 2016. Associations between hypothalamic-pituitary-adrenal axis function and peak bone mass at 20 years of age in a birth cohort. Bone 85, 37-44.