Published in final edited form as:

Psychosom Med. 2017 January ; 79(1): 119–121. doi:10.1097/PSY.0000000000000416.

Pursuing the Optimal Operationalization of Allostatic Load

Joshua F. Wiley, Ph.D.¹, Tara L. Gruenewald, Ph.D.², Arun S. Karlamangla, M.D., Ph.D.³, and Teresa E. Seeman, Ph.D.³

¹School of Psychological Sciences and Monash Institute of Cognitive and Clinical Neurosciences, Monash University, Victoria, Australia

²Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA

³Division of Geriatrics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA

In their commentary on our paper (1), Crook and Booth raise important questions regarding the theoretical and methodological appropriateness of applying factor analysis to model allostatic load (AL; 2). They argue factor scores not to be the ‘optimal’ measure of AL and conclude that more research is needed.

From a methodological perspective, Crook and Booth argue that the poor fit of the hierarchical model may be due to the proportionality constraint inherent in its structure. Whether the proportionality constraint or other features of the model are the sources, it does not change the result that the hierarchical model provided an inferior fit to the data than did the bifactor model. Considering the bifactor model, Crook and Booth suggested a stronger test of exchangeability by computing and correlating AL scores from non-overlapping biomarkers (2). However, unlike intelligence research, studies on biomarkers tend to have relatively few indicators as a consequence of feasibility factors (e.g., available blood sample volume, participant burden, cost). Separating already limited biomarker panels into two non-overlapping sets will be too few to estimate the bifactor model in many studies. As advances in multiplexing reduce barriers to assaying numerous biomarkers (3), we look forward studies that address this question.

Crook and Booth also note that, on average, the general factor explains about 11% of the variance in the biomarkers (2). In psychometric studies of carefully designed scales, 11% may be considered a small amount of variance, but this is expected for biomarkers for several reasons. First, except for the heart rate variability measures, each biomarker is distinct (e.g., interleukin 6 and tumor necrosis factor α are separate analytes with unique roles in the immune system and inflammatory processes). Second, biomarkers have circadian rhythms, which introduces variability due to the timing of assessment. Third, the timeframe reflected in each biomarker varies. For example, glycosylated hemoglobin approximately indexes the previous three months, overnight urinary epinephrine and

Address for Correspondence: Joshua F. Wiley, School of Psychological Sciences, 18 Innovation Walk, Monash University, VIC, 3800, Australia, T: +61 3 990 59598, joshua.wiley@monash.edu.

Conflicts of Interest and Sources of Funding
The author declares no conflicts of interest.

norepinephrine index approximately 12 hours, and blood pressure is comparatively momentary. Fourth, the general factor accounted for variance over and above the covariates age and sex. Therefore, we expected the general factor to account for a modest amount of the variance in biomarkers. Furthermore, considered over and above the effects of age, 11% of the variance is not necessarily trivial. Finally, the overall model (i.e., general + system factors and covariates) accounted for an average of 55% of the variance in biomarkers, rising to 60% when excluding soluble intracellular adhesion molecule 1 and low-density lipoprotein. If researchers believe including additional variability from individual biomarkers yields a more 'optimal' measure, principal component analysis (PCA) can be used (e.g., 4, 5). Using PCA, researchers can choose the tradeoff between dimension reduction and the desired amount of variance explained by extracting more components.

From a theoretical perspective, Crook and Booth argue the complex interrelations among biomarkers to be inconsistent with the model indicated by a bifactor model (2). We agree that the relations among biological measures are complex. Indeed, there is ample evidence that these biomarkers and systems relate to each other and dynamically to themselves via positive and negative feedback loops (e.g., 6). Crook and Booth argue that not accounting for these interconnected pathways biases results to favor the bifactor model (2). It is certainly implausible that one latent factor causes all of the biomarkers measured in our study. However, the practical implications of whether the general factor from the bifactor model captures variability in biomarkers due to a shared latent factor or instead captures variability due to numerous and bidirectional pathways among biomarkers are less clear. In either case, the general factor indexes a phenotype that emerges across multiple systems.

Nevertheless, from a pragmatic perspective, what alternatives are available? One possibility is graphical models, such as complex network analysis (7). If there are no latent or selection variables, then inference regarding possible causal relations can be extracted from observational data using do-calculus (8). As with many statistical models, the challenge lies in meeting the assumptions—no latent or selection variables. These assumptions are violated unless all variables related to assignment are known, measured, and included, which is typically only plausible with random assignment.

Rather than speculate that graphical models or network analysis may be more appropriate than a factor model, we tested this approach in the MIDUS data. Specifically, we extracted residual scores for each biomarker adjusting for age and sex and calculated the correlation matrix for all biomarkers using full information maximum likelihood estimation. These correlations were then subjected to the PC algorithm (7) to explore *potential* underlying causal structures that *could* give rise to the data. The results are shown using a graphical model (9) in Supplementary Figure 1. The potential underlying structure identified had biomarkers from related systems clustered together, although this is not surprising. Interestingly, the potential structure included cortisol, dehydroepiandrosterone sulfate, epinephrine, and norepinephrine as drivers of other biomarkers, but not vice versa. In comparison, although insulin resistance was upstream of triglycerides, waist-to-hip ratio, and E-selectin, it was also downstream from cortisol, epinephrine, and c-reactive protein. Thus the potential structure identified is strikingly consistent with AL theory on primary versus secondary mediators (10). It is worth noting that several closely-related measures of

parasympathetic nervous system activity were included in this analysis (high and low frequency heart rate variability as well as the standard deviation and root mean square of successive differences of the interbeat intervals). Highly correlated variables do not pose analytical challenges for factor analyses and the graphical models employed here, but would not be advisable to include simultaneously in some statistical models, such as regression analyses where multicollinearity is a concern.

Although graphical models can model complex interrelations among biomarkers and systems, without strict assumptions they cannot identify causal structures and they raise unique challenges. How would one operationalize a primary endpoint for a study or trial from such a complex network? Must 23 biomarkers be individually included to control for physiological dysregulation, and what consequences will this have for studies with modest sample sizes? In summary, we believe that the bifactor model of AL offers several practical advantages over alternate methods. When reporting results from latent factor scores, including supplementary analyses showing the results with individual biomarkers can help to clarify whether results are similar or vary across individual biomarkers, though caution is warranted as examination of individual results for many biomarkers will substantially increase the chance of Type 1 errors.

Biobehavioral research in psychosomatic medicine and related disciplines will benefit from further critical debate on research on novel measures of AL as well as statistical methods to quantify AL that appropriately capture our evolving understanding of the basic biology. Future methodological research that compares the strengths and weaknesses of alternate measures of AL in different contexts (e.g., small and large sample sizes, few or many biomarkers) would be especially useful to advance the appropriate use of AL. However, the greatest advancements in the pursuit of the optimal measure of AL may come from a deeper understanding of the processes that may give rise to AL, such as suggested in the letter by Picard and colleagues (11), more sophisticated assessments (e.g., stress reactive and functional measures of biomarkers; 12), functional measures of biomarkers), and more sophisticated designs (e.g., longitudinal and experimental studies).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Wiley JF, Gruenewald TL, Karlamangla A, Seeman TE. Modeling multisystem physiological dysregulation. *Psychosom Med*. 2016;78:290–301. [PubMed: 26734956]
2. Crook Z, Booth T. Considering the Appropriateness of the Factor Analytic Operationalization of Allostatic Load. *Psychosom Med*. in press.
3. Mills PJ, Peterson CT. Multiplexing and Beyond in Biobehavioral Research. *Psychosom Med*. 2016.
4. Cohen AA, Li Q, Milot E, Leroux M, Faucher S, Morissette-Thomas V, Legault V, Fried LP, Ferrucci L. Statistical distance as a measure of physiological dysregulation is largely robust to variation in its biomarker composition. *PloS one*. 2015;10:e0122541. [PubMed: 25875923]
5. Wiley JF, Carrington MJ. A metabolic syndrome severity score: A tool to quantify cardio-metabolic risk factors. *Prev Med*. 2016;88:189–95. [PubMed: 27095322]

6. Biddinger SB, Kahn CR. From mice to men: insights into the insulin resistance syndromes. *Annu Rev Physiol.* 2006;68:123–58. [PubMed: 16460269]
7. Kalisch M, Mächler M, Colombo D, Maathuis MH, Bühlmann P. Causal inference using graphical models with the R package pcalg. *J Stat Softw.* 2012;47:1–26.
8. Pearl J *Causality: models, reasoning and inference.* 2nd ed: Cambridge University Press; 2009.
9. Epskamp S, Cramer AO, Waldorp LJ, Schmittmann VD, Borsboom D. qgraph: Network visualizations of relationships in psychometric data. *J Stat Softw.* 2012;48:1–18.
10. McEwen BS, Seeman T. Protective and damaging effects of mediators of stress: elaborating and testing the concepts of allostasis and allostatic load. *Ann N Y Acad Sci.* 1999;896:30–47. [PubMed: 10681886]
11. Picard M, Juster R-P, Sloan RP, McEwen BS. Mitochondrial Nexus To Allostatic Load Biomarkers. *Psychosom Med.* in press.
12. Juster R-P, Moskowitz DS, Lavoie J, D'Antono B. Sex-specific interaction effects of age, occupational status, and workplace stress on psychiatric symptoms and allostatic load among healthy Montreal workers. *Stress.* 2013;16:616–29. [PubMed: 23952366]