# University of Groningen

## Are we witnessing the decline effect in the Type D personality literature? What can be learned?

Coyne, James C.; de Voogd, Jacob N.

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*
Publisher's PDF, also known as Version of record

Link to publication in University of Groningen/UMCG research database

Correspondence

# Are we witnessing the decline effect in the Type D personality literature? What can be learned?

James C. Coyne [a,b,*], Jacob N. de Voogd [a,c]

[a] Health Psychology Section, Department of Health Sciences, University Medical Center Groningen, University of Groningen, The Netherlands
[b] Department of Psychiatry, Perelman School of Medicine of the University of Pennsylvania, USA
[c] Department of Pulmonary Rehabilitation, Center for Rehabilitation, University Medical Center Groningen, The Netherlands

### ARTICLE INFO

### ABSTRACT

After an unbroken series of positive, but underpowered studies seemed to demonstrate Type D personality predicting mortality in cardiovascular disease patients, initial claims now appear at least exaggerated and probably false. Larger studies with consistently null findings are accumulating. Conceptual, methodological, and statistical issues can be raised concerning the construction of Type D personality as a categorical variable, whether Type D is sufficiently distinct from other negative affect variables, and if it could be plausibly assumed to predict mortality independent of depressive symptoms and known biomedical factors, including disease severity. The existing literature concerning negative affect and health suggests a low likelihood of discovering a new negative affect variable that independently predicts mortality better than its many rivals. The apparent decline effect in the Type D literature is discussed in terms of the need to reduce the persistence of false positive findings in the psychosomatic medicine literature, even while preserving a context allowing risk-taking and discovery. Recommendations include greater transparency concerning research design and analytic strategy; insistence on replication with larger samples before accepting "discoveries" from small samples; reduced confirmatory bias; and availability of all relevant data. Such changes would take time to implement, face practical difficulties, and run counter to established practices. An interim solution is for readers to maintain a sense of pre-discovery probabilities, to be sensitized to the pervasiveness of the decline effect, and to be skeptical of claims based on findings reaching significance in small-scale studies that have not been independently replicated.

© 2012 Elsevier Inc. All rights reserved.

## Introduction

Many "discoveries" turn out to be exaggerated or simply false across diverse research areas and disciplines. John Ioannidis [1] provoked concern with his demonstration that of the 49 most cited clinical research studies in major journals, replications of 34 had been attempted, and 41% of key findings had been refuted or shown to be substantially diminished. In a subsequent paper [2] he offered evidence that "most claimed research findings are false" and he and others [3–6] have since identified some mechanisms by which discoveries appear in the literature and then undergo a decline or refutation. Claims of breakthrough discoveries frequently arise in small positive studies that would seem to be too underpowered to detect an effect, and the inadequate sample size makes findings all the more attention grabbing.

Yet, portrayals in the literature of such "discoveries" ignore the larger context of a strong confirmatory bias in published research papers, with unknown numbers of negative findings not being published and findings being declared discoveries simply because of having achieved an arbitrary level of significance. Moreover, statistically significant findings in small studies are necessarily always large because larger effect sizes are needed to cross the higher threshold for significance. Apparent discoveries are thus created and perpetuated by a combination of confirmatory bias, flexible rules of design, data analysis and reporting [5,6], and significance chasing [3]. There is a recurring decline effect [4] in diverse literatures, usually occurring when discoveries declared on the basis of small studies attract more resources or the attention of new investigator groups with less of an investment in the discovery, and attempted replications or extension of the original findings fail.

## The rise and apparent decline of Type D personality

Are we now witnessing a decline effect in the literature concerning Type D personality? The concept, which has been defined as the tendency to experience negative emotions and to inhibit self-expression in social interaction, was riding high as a promising prognostic indicator for mortality in cardiovascular disease (CVD) patients. A preliminary

\* Corresponding author at: Department of Psychiatry, Perelman School of Medicine of the University of Pennsylvania, 3535 Market St. Rm 676, Philadelphia, PA 19104, USA. Tel.: +1 215 662 7035; fax: +1 215 349 5067.
E-mail address: jcoyne@mail.med.upenn.edu (J.C. Coyne).

study of post myocardial infarction (MI) patients had employed a different distressed personality type constructed by cross tabulating dichotomized continuous measures of trait anxiety and social inhibition, selecting the high/high quadrant and compared it to the other three [7]. Low exercise tolerance, a previous MI, anterior site of the MI, smoking, and age were entered into a logistic regression predicting 15 deaths and it was found that addition of the distressed personality type added to the prediction. A subsequent *Lancet* study constructed a Type D personality with different, but similar personality measures and claimed that patients with Type D personality were four times more likely be among the 21 dead during the observation period [8]. A later study claimed that patients with Type D personality were ten times more likely to die after a cardiac transplant [9]. Claims were also made that its predictive value was independent of disease severity and other biological variables. An unbroken series of five published studies [8–12] found a significant relationship between Type D personality and subsequent death.

However, all studies were conducted by the same investigator group at Tilburg University, with very small numbers of deaths being explained in each study: 15 in the preliminary study [7], 21 in the *Lancet* study [8], followed by 6 [10], 4 [11], 6 [9], and 12 [12] respectively (See Table 1). These studies consistently found Type D to be associated with mortality, but none of the studies had enough deaths to reliably detect a significant predictor of mortality with a strength similar to other predictors of mortality in heart disease, even if the effect was present. Observed magnitudes of effect were excessive relative to what could be reasonably expected and, in most cases, far stronger than traditional risk factors for mortality and recurrence, unless an extraordinary and unprecedented psychological risk for mortality had been uncovered.

After the initial "discovery" of an effect on mortality in a sample with as few deaths to explain as was the case in Type D studies, the continued reliance on small studies implicitly committed investigators to the assumption that they were pursuing a replication of an effect that is larger than many traditional biomedical risk factors. In hindsight, this commitment should have been made explicit and should have required elaborate justification beyond the initial finding with a small sample.

Moreover, the statistical significance of results in these earlier studies was typically vulnerable to reclassification or addition or subtraction of a few or even a single death. Observed results might well be due to capitalization on chance or confirmatory bias allowed by such factors as flexible extension or contraction of follow-up periods; selective exclusion of some deaths; selective definition of endpoints (cardiac-specific mortality versus all cause mortality versus mortality included in a composite endpoint); selective inclusion of possible covariates and exclusion of others, depending on the impact on Type D significance and magnitude; or reporting of significance levels only for adjusted, rather than unadjusted analyses. For example, reports of these studies offered no indication that follow-up periods and their stopping points were predetermined or whether follow-up was stopped or extended, depending on whether a significant prediction of mortality had been obtained.

Subsequently, the Tilburg group produced three studies that had more deaths to explain and therefore a greater likelihood of reliably detecting the value of Type D as a prognostic indicator for mortality. The first of the investigator groups' larger studies attempted to predict 47 deaths, but the prognostic value of Type D personality did not hold, once appropriate controls were introduced [13]. The second study attempted to predict 123 deaths among 641 heart failure patients, but Type D personality failed to be a significant prognostic indicator in either bivariate or multivariate analyses with statistical controls [14]. The third study attempted to predict 187 deaths among 1234 consecutive CVD patients receiving a percutaneous coronary intervention and found no association between Type D personality and mortality [15].

Recently, there have been several studies from outside the original investigator group. The first one failed to find a prognostic value for Type D, but was underpowered to do so, with only 11 deaths being explained [16]. Then, another two more studies from outside the original investigator group were published, with larger numbers of

**Table 1**
Studies assessing Type D and mortality in patients with cardiovascular diseases

| Publication | Sample size | Number of deaths | Type D measure | Type D prevalence | Unadjusted effect size with confidence interval | Adjusted effect size with confidence interval | Number of covariates |
|---|---|---|---|---|---|---|---|
| Denollet 1995 [7] | 105 | 15 | SI: heart patients psychological questionnaire ; NA: trait STAI median split SI≥12, NA≥40 | 27% | *nr* | *nr* | 5 |
| Denollet 1996 [8] | 210 | 21 | SI: heart patients psychological questionnaire; NA: trait STAI; median split SI≥12, NA≥43 | 29% | *nr* | OR 4.1 CI 1.9–8.8 | 5 |
| Denollet 2000 [10] | 319 | 6 | DS16; median split NA≥9, SI≥15 | 31% | OR 11.65 CI 1.34–101.07 | OR 8.9 CI 3.2–24.7 | 5 |
| Denollet 2006 [11] | 337 | 4 | DS16; median split A≥9, SI≥15 | 29% | OR 4.84 CI 1.42–16.52 | *nr* | *nr* |
| Pedersen 2007 [12] | 358 | 12 | DS14; NA/SI≥10 | 30% | OR 2.51 CI 1.09–5.82 | HR 2.61 CI 1.12–6.09 | 9 |
| Denollet 2007 [9] | 51 | 6 | DS14: NA/SI≥10 | 29% | OR 16.5 CI 1.72–158.22 | HR 11.33 CI 1.24–103.3 | 2 |
| Schiffer 2010 [13] | 232 | 47 | DS14; NA/SI≥10 | 21% | OR 2.16 CI 1.05–4.43 | HR 1.40 CI 0.93–4.29 | 3 |
| Pelle 2010 [14] | 641 | 123 | DS14; NA/SI≥10 | 20% | OR 1.16 CI 0.72–1.87 | HR 1.09 CI 0.67–1.77 | 21 |
| Volz 2011 [16] | 111 | 11 | DS14; NA/SI≥10 | 30% | OR 1.40 CI 0.38–5.14 | HR 0.91 CI 0.25–3.32 | 2 |
| Grande 2011 [17] | 977 | 172 | DS14; NA/SI≥10 | 25% | OR 0.84 CI 0.57–1.25 | HR 0.99 CI 0.61–1.59 | 19 |
| Coyne 2011 [18] | 706 | 192 | DS14; NA/SI≥10 | 13% | HR 0.89 CI 0.58–1.37 | HR 0.78 CI 0.49–1.24 | 2 |
| Damen 2012 [15] | 1234 | 187 | DS14; NA/SI≥10 | 29% | *nr* | HR 1.19 CI 0.76–1.85 | 22 |

SI: social inhibition, NA: negative affectivity, STAI: State-Trait Anxiety Inventory, DS16: Type D scale 16-item version, DS14: Type D scale 14-item version, CI: confidence interval, *nr*: not reported in original paper.
N.B.: length of follow-up is not reported because descriptive statistics were inconsistent.

deaths to be explained than most of the previous studies, but neither found evidence for a prognostic value of Type D personality in unadjusted or adjusted models, despite one having 172 deaths among 977 patients with CVD [17] and the other study having 192 deaths at 18 months follow-up of 706 HF patients [18].

Both of the studies relied on all cause, rather than cardiac-specific mortality as the primary outcome. The authors could justify the reliance on all cause mortality since cause of death is not reliably reported in death certificates and because of the unreliability of judgments that death is cardiac specific in an elderly population with numerous comorbidities who do not die at a research hospital. One of these studies [18] had a shorter follow-up period (18 months) than earlier studies and a lower prevalence of Type D personality. However, there was a nonsignificant trend in the wrong direction in terms of prediction of mortality from Type D personality and a large enough number of deaths being explained so that the possibility can be dismissed of the trend likely being reversed with continued follow-up.

More importantly, this study [18] provided a critique of the methods used to identify Type D patients in previous studies and the overfitted regression models used in testing multivariate prediction controlling for potential confounders [19]. These factors would be expected to lead to spurious findings of a prognostic value for Type D. These points were echoed and amplified in an accompanying editorial [20].

## Is a personality type tenable conceptually and statistically?

The Tilburg group identified patients as having a Type D personality if they were above the median [8,10] or, in later studies [9,13,14], if they scored above 10 on continuous measures of both negative affectivity (NA) and social inhibition (SI), and then compared the outcomes of these patients in the high/high quadrant to the other three quadrants. Such an analytic strategy for combining continuous predictor variables has been widely rejected for decades in the personality, education, and basic statistics literatures [21–27] because it likely leads to spurious results. There are two objections: splitting continuous variables into arbitrarily dichotomized categories, but also the isolation of the high/high category as a personality type.

Whereas dichotomizing a single continuous variable loses information and statistical power, cross tabulation of a pair of dichotomized variables is likely to lead to inflated estimates of statistical significance and dramatically increased risk of spurious findings [24]. Humphreys [27] declared the construction of a $2 \times 2$ matrix from continuous variables to be "unnecessary, crude, and misleading" (p. 874) and Cohen and Cohen [26] called it "an abuse of data" (p. 310). In labeling a person as an introvert or as neurotic, theorists adopted a typological language as "a verbal convenience rather than a meaningful mode of categorization" (p. 1159) [28] without any evidence that a specific demarcation point exists, at which persons above this point not only resemble each other, but differ from persons below it in crucial ways [21].

While the success of typological thinking in biology can be demonstrated in the classification of plants and animals, examples in psychology are harder to identify. Thus, biological sex as male or female approaches a sharply defined category, but the representation of the psychological dimensions of masculinity and femininity does not [28]. Applied to the assessment of Type D personality, the implications are for a strong preference for preserving NA and SI as continuous variables and examining their interaction, rather than postulating a sharp distinction between high/high quadrant and the other three quadrants in a dichotomization of these variables. Among the problems in isolating the high NA/high SI quadrant to construct Type D personality is that NA and SI are moderately correlated and so patients are selected for greater distress than if either of the measures were considered separately. Scoring high on two correlated measures of distress is a more reliable indication of distress than being high on only one.

Long-standing concerns about statistical and conceptual difficulties of categorical/typological constructs have been substantiated in recent work using sophisticated simulated comparison data techniques [29,30] that consistently favors dimensional conceptualizations. One systematic review [29] compared 377 articles with 311 distinct findings and concluded:

> The domains of normal personality, mood disorders, anxiety disorders, eating disorders, externalizing disorders, and personality disorders (PDs) other than schizotypal yielded little persuasive evidence of taxa [categories].... This review indicates that mostly variables of interest to psychiatrists and personality and clinical psychologists are dimensional, and that many influential taxonic findings of earlier taxometric research are likely to be spurious (p. 903).

Ferguson and colleagues [31] applied two taxometric procedures MAMBAC and MAXCOV to scores from the DS14, the standard measure of Type D personality and found clear evidence for a representation of it in dimensional rather categorical terms. The strong suggestion of these findings is that Type D research should refocus on the additive and multiplicative influence of NA and SI in the context of other risk factors. While it is possible that inflection points for the continuous variables could be identified that maximize influence of these variables on particular CVD outcomes, the task remains of determining whether such inflection points need to be identified for NA and SI separately or whether there should be single or sliding cutpoints specified for the interaction term with different outcomes.

## Rescuing Type D personality as an interaction term in a regression analysis of continuous variables?

There are multiple reasons for doubting the validity of earlier claims from small studies that categorical Type D personality defined by the high/high quadrant predicted mortality. The two larger studies conducted outside the Tilburg group [17,18] had analyzed Type D personality data using both the scoring procedure of the original investigator group, but also by preserving NA and SI as continuous variables and examining their interaction effect, and in neither instance found a significant bivariate or multivariate prediction of mortality. Could claims about Type D personality's independent prognostic value nonetheless be revived by reanalysis of the data from the earlier studies that preserved the continuous nature of component variables? Probably not. It is unlikely that a significant interaction effect would support categorical over dimensional conceptions of Type D personality or that the optimal weighting of NA and SI would correspond to the arbitrarily chosen scoring of Type D personality in past studies. More basically, given the small number of deaths being explained in the positive studies, it is highly unlikely that the interaction term would prove significant. Smith [20] pointed out that even if NA or SI alone or some additive combination predicted mortality, there was the risk of what Block [32] has termed the "jangle fallacy" of assuming that simply assigning a new name to previously studied traits constitute a discovery.

However, Smith [20] provided a more profound criticism of relying on the one cell versus three (high NA and high SI versus the three other combinations of these variables) to test the specific prediction of a synergistic effect of NA and SI. The three versus one contrast could occur for numerous other reasons than a synergy of these two variables. Namely, a significant contrast could be produced by an additive effect of either NA or SI alone; two additive first order effects without any trace of an interaction. We add that a spurious interaction effect due to either an overlap in unmeasured nonlinear terms [33,34], or, as we have noted, the benefits of combining two imperfectly reliable measures of negative affect rather than relying on one imperfectly reliable measure.

## Meta-analysis to the rescue?

Could the entire body of studies relating Type D personality to mortality be combined in a meta-analysis? Grande and colleagues

[35] performed such a meta-analysis of available studies and the resulting forest plot (Fig. 5 in [35]), offered a striking graphic representation of what was already known. The earlier smaller studies by the Tilburg group were positive and the two later, larger scale studies conducted outside the group are negative. With understatement, the authors of the meta-analysis declared that "recent methodologically sound studies suggest that early Type D studies had overestimated the prognostic relevance (p. 299)" and ended with a call for "urgently required," more sophisticated studies. But these conclusions take for granted that there is a Type D personality phenomenon to be explained in the absence of a single positive finding having any of the following characteristics: (a) appropriate analyses of data from (b) an adequately powered study (c) conducted outside the original investigator group.

This meta-analysis relied on the Tilburg group's categorization of a Type D personality, rather than a regression analysis of continuous variables and their interaction. The overall effect size was significant, which simply attests to the ability of a substantial portion of small underpowered studies to dominate a meta-analysis. Moreover, the considerable heterogeneity between studies begged explanation, without which the interpretation of the overall effect size would not be meaningful.

Responding to criticism [36] of this meta-analysis, Grande [37] proposed a patient level meta-analysis as a solution. This would require obtaining original continuous NA and SI data from the authors of previous studies, but for two reasons would be unsatisfactory solution and would only serve to obscure some obvious heterogeneity and bias in the contributing studies [38]. First, the validity of such a patient level meta-analysis assumes that all available data are included without substantial bias. Our group can attest to a publication bias in studies of Type D personality, even if we cannot quantify it. We had difficulty publishing our large study of Type D personality predicting mortality [18], with it getting rejected from both *Lancet* and a number of cardiology journals that published positive studies of Type D personality predicting other cardiac outcomes. It is even less likely that negative findings from unpowered studies would be published because of confirmatory bias and the further justification for rejection that such studies were too small to expect a significant finding. Second, positive studies have differed in key features such as beginning and end of observation periods, means of assessing Type D personality, and availability and selection of covariates. Unless we are willing to ignore such considerations, meaningful integration becomes difficult or impossible. Finally, the consistent series of published underpowered studies yielding strong positive effects involved unknown prior analyses leading to decisions about beginning and ending of follow-up periods and selection and coding of covariates. Think of it: would we accept the validity of a calculation of odds based only on betting sequences that always ended with winnings and with no losing sequences included, despite the prior expectation that odds were favorable to the house?

Overall, obtaining complete, unbiased patient-level data for a meta-analysis may be more difficult when the data are derived from observational, correlational studies of the prediction of mortality than when the data represent the declared outcomes of clinical trials. The existence of unpublished clinical trial data may be more easily detected than the existence of aborted examinations of mortality data associated with observational studies, particularly when baseline data for the study were collected for other purposes and the follow-up assessment points were not fixed. There may be no record of multiple examinations of data with adjustments in the start and ending of the follow up.

## Type D personality predicting mortality independent of depression and known biomedical prognostic factors: déjà vu all over again?

Even acknowledging conceptual and statistical criticisms and the possibility of a jangle fallacy, claims remain provocative that Type D personality predicts mortality independent of depression and biomedical prognostic variables including severity of disease. These claims might seem for some to warrant further investigation before dismissing the validity of some sort of reformulated Type D personality.

First, these claims arise in multivariate regressions with the number of covariates included in models are a substantial proportion of the number of deaths being explained. Rules of thumb for minimal covariates/events vary, but most authorities agree a minimum of 10 to 15 events are needed per covariates to avoid generating spurious associations [39]. Undoubtedly even those who advocate relaxation of these rules would balk at having as high a ratio of covariates to number of events being explained.

Furthermore, claims that prognostic value of Type D personality is independent of depressive symptoms should be greeted with skepticism, given the long-recognized high association among measures of negative affect [40–44]. Seemingly reasonable theoretical and conceptual distinctions often cannot be sustained when attempts are made to demonstrate discriminant validity of measures of specific negative affects.

The NA component of Type D personality is as highly correlated with self-reported measures of depressive symptoms as their respective internal consistencies allow [18]. Any apparent advantage in prognostic value of Type D personality versus depression is an artifact of the construction of the high/high typology and it should be expected that substituting depressive symptoms for NA would generate similar spurious advantages for a reformulated Type D personality. Finally, with measures of negative affect so highly correlated, it is not clear why a further substantial contribution to predicting mortality would be expected from adding consideration of depressive symptoms. Whatever would be represented by Type D personality controlling for depressive symptoms would have little resemblance to what is intended by the unadjusted Type D personality variable [20]. Statistical adjustments of highly correlated negative affect variables can generate anomalies as seen in a finding that anxiety is inversely related to mortality after other negative affect variables entered into multiple regression analyses, and yet anxiety can be shown to be an independent positive predictor when considered alone [45]. In short, there is a large literature relating various measures of negative affect to CVD outcomes that provides little basis for the expectation that singling out a combination of NA and SI will have much consistent advantage over other such negative affect variables, individually or combined.

As for the claim that Type D personality predicts death independent of established biomedical factors, it too arises in overfitted multiple regression equations. Moreover, there is a large literature showing that negative affect, and particularly depressive symptoms can often be found to predict mortality, but this literature also indicates considerable difficulty singling out particular psychological variables as a risk for mortality or that the association is not artifactual. Ketterer and colleagues [40] have referred to the "big mush" of confounded and non-independent negative affect measures. Meehl [46] has applied the label "crud factor" to the broader tendency of self-reported negative factors to be correlated with each other in ways that cannot readily be unambiguously differentiated. Ormel and colleagues [42] have suggested that NA/neuroticism has little explanatory value because it is related to so many diverse self-reported negative antecedents, concurrent conditions and consequences. He noted evidence that NA/neuroticism involves a negative response style that extends to other self-reports. While it might seem that a hard outcome like mortality would escape the criticism of confounding and response style, MacLeod and colleagues [47] cogently argue that the challenge is to distinguish the causal influence of negative affect from other negative environmental and physical health variables. They have provided a number of demonstrations of the anomalies and inconsistencies that can arise when drawing causal inferences about apparent associations of self-reported depressive symptoms or stress with physical health outcomes. There is a high likelihood of noncausal

relationships generated by confounding between self-reported negative affect and physical health outcomes, with residual confounding often proving impossible to discount. What needs to be ruled out – and generally cannot be – is that the negative affect variable merely reflects unmeasured or imperfectly measured aspects of overall disease status, including comorbidity.

Plausible putative psychophysiological mechanisms can be invoked to explain particular negative affects being related to mortality, but these mechanisms often work just as well for rival negative affect variables or for a relationship found to be in the opposite direction [48]. The task of disambiguating any independent causal role of the negative affect in death is shifted to the similarly difficult task of establishing the variable's role in a poorly mapped, complex psychophysiological process in which the precise role of easily measured physiological variables cannot yet be identified. Showing that a particular negative variable is correlated with some physiological variable is easier than establishing that it mediates the causal influence of the negative affect variable in a particular psychophysiological process [49].

In general, the efficacy of introducing physical health covariates has been overestimated as a means to rule out confounding physical health variables as the source of an association between self-reported negative affect and physical health outcomes in observational studies of aged chronically ill patient samples. "Statistical adjustment using linear component of imperfect measure of disease severity is not sufficient to discount such a possibility" [50]. This applies not only to disease severity, but to typically overlooked physical comorbidity and overall disease burden. A simple bivariate association of negative affect with physical health outcomes is ambiguous and could be explained by confounding with correlated health variables. Yet, application of statistical controls often produces a more biased and confounded estimate of the influence of negative affect on physical health outcomes. The efficacy of statistical controls depends upon the unlikely achievement of all of the relevant potential confounders having been identified, temporal ordering having been established, mediators having been distinguished from confounders, and all relevant variables having been measured with precision.

Standard practices such as introducing all covariates into regression equations, pre-selecting them on the basis of correlations with the outcome, or relying on automatic selection procedures such as backwards regression all can generate considerable proportion of noise variables [51], and – particularly when there is even moderate correlation among the disease variables – exclude potentially important independent risk factors [52]. "Statistical adjustment by an excessive number of variables or parameters, uninformed by substantive knowledge (e.g. lacking coherence with biologic, clinical, epidemiological, or social knowledge)…can obscure a true effect or create an apparent effect when none exists."[53]. What is needed and has not been available in studies of Type D personality predicting mortality is a precise model specification. Presumed "replications" of positive findings need to adhere to the same model or explain why another model was used, but the usual practice has been to declare a replication anytime the positive results are found for Type D personality.

## Does death matter anymore in psychosocial research predicting CVD endpoints?

While numerous other investigator groups have been exploring the associations among Type D and other self-report negative affect and health-related variables, the cachet of these studies depends on claims from the Tilburg studies that Type D represents a unique, strong risk factor for death. Yet, the practical clinical implications of any association between Type D personality and mortality in CVD may not be as clear as has been assumed. Type D personality has been construed as a relatively immutable trait and so there would no obvious move to intervention studies to modify Type D personality. A recent trial of a multimodal intervention failed to modify Typed

D personality [54]. Perhaps patients with Type D personality could be targeted for interventions that improve their disease specific health status, but it has not been clear why it would not be more efficient simply to target relevant physical health variables, unmoderated by Type D personality. Calls for Type D personality being used for routine screening of CVD patients have not been grounded in studies of the performance characteristics of measures of Type D personality as screening instruments and the cutpoints that have been advocated have been arbitrary. It is not clear what the clinical application would have been of such screening: given the small proportion of deaths occurring in studies purporting to demonstrate a prediction of mortality, the optimal prediction strategy would have been to disregard the results of a screening with a Type D personality measure and predict that each individual patient will survive. And given the lack of evidence for clinically relevant predictive value, it would simply have been unethical to exclude patients from cardiac rehabilitation solely on the basis of their Type D scores. The main appeal of Type D personality seems to have been the claim that it identified an independent predictor and therefore risk factor for mortality.

Claims that psychological variables can predict mortality and that their modification can extend life of ill and dying people are important to the credibility and prestige of psychosomatic medicine, and there may be a greater willingness to entertain claims that psychological variables are related to mortality and even to protect these claims from disconfirmation. Much can be learned from the literature concerning depression in cardiac patients. Depression is a psychological variable that has been given particular attention in behavioral cardiology, not only because of studies showing an association with mortality, but because it is modifiable. Yet, a systematic review found a lack of any studies indicating that routine screening for depression improves cardiac patients' physical health outcomes [55]. Moreover, issues have been raised about the practical clinical relevance of mortality as an outcome. Advances in medical treatment of cardiac conditions [56] now requires as enrollment of as many as 14,000 patients to demonstrate that a clinical procedure or device has an advantage over the patient survival achieved in routine care. So, we are left with the difficulties establishing causality in observational studies and the likely impracticality of demonstrating causality in randomized controlled trials for Type D personality as well as depression.

## Summary: can we save the credibility of psychosomatic medicine from repeated decline effects?

The literature concerning Type D predicting mortality personality seems to fit the pattern of a discovery in a seemingly underpowered study subsequently being found to be a false positive. Not only is there a series of larger null trials accumulating in the literature, fundamental issues have been raised about the validity of the conceptualization of Type D personality as a categorical type. We can expect that further studies of Type D personality will be required to analyze personality data in terms of the interaction term of continuous variables or offer a compelling reason why the standard strategy should continue to be used. Also, as we have seen here, further issues can be raised when Type D personality is considered in the context of the larger literatures relating negative affect variables to health. Once the distinctiveness of Type D is questioned, it becomes easier to identify difficulties progressing to an identification of Type D as a causal factor in CVD outcomes, particularly given the precedent of the literature concerning depression. A commentary [57] that accompanied the original Type D mortality study [8] in *Lancet* in 1996 seems prescient:

> Denollet et al. added a new term – the distressed personality (Type D) – to a field congested with related concepts including type A personality, anger and hostility, psychological stress, vital exhaustion, major depression, depressive symptoms, and social

isolation. Each of these concepts enjoyed a period of prime time exposure following publication of one or more epidemiological reports linking it to mortality in patients with CHD and then declined in popularity…. (p. 414)

Claims that Type D personality predicted mortality attracted considerable attention, even being featured in a story in *Time* magazine [58]. Such claims are newsworthy and add to the short-term credibility and prestige of psychosomatic medicine. Yet, there is the risk that a recurring pattern of a particular association declining with the association being shown to be spurious, exaggerated, or simply false, will undermine the credibility of the field. Previous claims that were shown to be exaggerated or simply false contributed to the decline and almost demise of psychosomatic medicine in the 1950s and 1960s [59].

It would be beneficial to reflect on the conditions that led to a series of published underpowered studies seemingly replicating the initial finding of Type D personality predicting mortality. Why were substantially underpowered studies repeatedly accepted for publication after the initial discovery study, rather than editors and reviewers insisting on larger number of deaths being explained? Why were there no challenges to postulating of a personality type when there was such a weight of evidence that dimensions were superior to categorical conceptualizations of personality, and the data reduction analyses techniques used to construct categories were prone to spurious results? Why was the observation, made at the time of the first Type D personality–mortality study not followed up: that Type D personality was entering into already congested area of seemingly similar concepts [57]? Why was there apparently no insistence from peer reviewers that the investigators address these issues as a condition for publication or and why were these criticisms not expressed in post-publication commentary like letters to the editor? The answers to these questions may point to the proneness of the field to confirmatory bias, particularly in claims about psychological factors influencing mortality, and therefore the field's vulnerability to future decline effects, as well as its limited ability to self correct in a timely fashion.

There is a need for reforms in the standards for evaluating claims as "discoveries" and for publication practices in psychosomatic medicine. True breakthrough discoveries can be expected to occur infrequently and even when claims of discoveries are not based on confirmatory bias and significance chasing, they can often be expected not to survive attempts at replication. Efforts to uncover important new clinically relevant phenomena are thus risky and prone to failure and so there is a need for not making the search more difficult than it already is. The discovery process can be expected to be "unfettered, haphazard, exploratory, opportunistic, selective, and highly subjectively interpreted" [60, p 645]. Although not all would agree, some advocates for reform suggests that flexible rules of data analysis and even fishing expeditions are still permissible, but only if all results are reported [6]. However, investigators should be disallowed from presenting their discoveries as if they were obtained by conventional confirmatory, rather than exploratory analyses, and the full range of methods and examinations of the data conducted before arriving at the "discovery" need to be identified transparently. Acceptance of a discovery should only be provisional until after independent replication with strict adherence to conventional confirmatory practices, and replication should entail obtaining precisely similar results, rather than simply achieving statistical significance.

Across fields, there have been a number of suggestions for changes in editorial practices that are intended to make the occurrence of decline effects less common [61–65]. These include some fundamental changes like publication of negative findings and non-replications; prior registration of the protocols for observational studies, much the same as clinical trials are now being registered; making publicly available data sets for reanalysis; and not accepting "discoveries" until they are independently replicated. A code of conduct has been suggested for investigators that would be enforceable by reviewers [6]. For observational studies predicting mortality, rules might include having minimal sample sizes, evidence of pre-specified follow up periods and analytic plans, disclosure of both simple bivariate and multivariate results, and demonstration that positive results are not tied to arbitrary analytic decisions.

Many of these suggestions are well intended, but run counter to established practices that continue to be rewarded. Many journals, and particularly high impact journals, do not publish null findings or even replication studies, because greater prestige is accorded initial discoveries. A survey suggests that psychologists openly admit multiple examinations of data before settling on the most positive findings and to suppressing negative findings [66]. A prominent psychologist has openly advocated that investigators change their stated hypotheses to what fits the results "where the data may be strong enough to justify recentering your article around the new findings and subordinating or even ignoring your original hypotheses." [67]. Overall, publishers journals, reviewers, funding agencies, department heads, and the media all reward extravagant claims and confirmatory bias and outright hype [68].

Many of the reforms that have been suggested to reduce the possibility that most discoveries will turn out to be false or exaggerated and a consequent inevitable decline effect are unlikely to be achieved and may be unenforceable. Journals continue to resist publishing attempted replications that fail or simply negative findings because of the greater prestige of positive discovery studies. It also would be difficult to enforce the requirement that investigators conducting observational studies formally register their hypotheses, length of follow-up, and covariates being considered before commencing their analyses.

Open-minded but skeptical readers and reviewers are left to their own devices in deciding whether new declarations of discoveries are credible or are likely undergo a decline effect. Relying on prior probabilities derived from specific research literatures but also more general observations about the pervasiveness of decline effects, they can ask for themselves in the case of a new discovery of a Type D personality-like phenomenon 'what is the likelihood of a distinct negative affect being singled out as an independent predictor of mortality superior to its rivals and with appropriate controls for biomedical risks factors?' They can be skeptical of discoveries emerging in underpowered trials that have not yet been independently replicated and they can be skeptical of "replications" that might have been achieved with repeated examinations of the data with flexible follow-up periods, exclusion criteria, covariates being considered and other design and analytic decisions. In other literatures and notably in genome wide studies, the tendency has been noted to consider particular findings replicated when any positive result was achieved, even if not a replication of the specific "discovery." Stricter criteria for declaring replications should be applied, if not by journal editors and reviewers, then by skeptical readers.

## References

[1] Ioannidis JPA. Contradicted and initially stronger effects in highly cited clinical research. JAMA 2005;294:218-28.
[2] Ioannidis JPA. Why most published research findings are false. PLoS Med 2005;2: 696-701.
[3] Ioannidis JPA. Evolution and translation of research findings: from bench to where. PLoS Clin Trials 2006;1:e36.
[4] Schooler J. Unpublished results hide the decline effect. Nature 2011;470:437.
[5] Simmons JP, Nelson LD, Simonsohn U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychol Sci 2011;22:1359-66.
[6] Wagenmakers EJ, Wetzels R, Borsboom D, van der Maas HL. Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). J Pers Soc Psychol 2011;100:426-32.
[7] Denollet J, Sys SU, Brutsaert DL. Personality and mortality after myocardial infarction. Psychosom Med 1995;57:582-91.
[8] Denollet J, Sys SU, Stroobant N, Rombouts H, Gillebert TC, Brutsaert DL. Personality as independent predictor of long-term mortality in patients with coronary heart disease. Lancet 1996;347:417-21.

[9] Denollet J, Holmes RVF, Vrints CJ, Conraads VM. Unfavorable outcome of heart transplantation in recipients with Type D personality. J Heart Lung Transplant 2007;26:152-8.

[10] Denollet J, Vaes J, Brutsaert DL. Inadequate response to treatment in coronary heart disease — adverse effects of Type D personality and younger age on 5-year prognosis and quality of life. Circulation 2000;102:630-5.

[11] Denollet J, Pedersen SS, Vrints CJ, Conraads VM. Usefulness of Type D personality in predicting five-year cardiac events above and beyond concurrent symptoms of stress in patients with coronary heart disease. Am J Cardiol 2006;97:970-3.

[12] Pedersen SS, Denollet J, Ong ATL, Sonnenschein K, Erdman RAM, Serruys PW, et al. Adverse clinical events in patients treated with sirolimus-eluting stents: the impact of Type D personality. Eur J Cardiovasc Prev Rehabil 2007;14:135-40.

[13] Schiffer AA, Smith ORF, Pedersen SS, Widdershoven JW, Denollet J. Type D personality and cardiac mortality in patients with chronic heart failure. Int J Cardiol 2010;142:230-5.

[14] Pelle AJ, Pedersen SS, Schiffer AA, Szabó B, Widdershoven JW, Denollet J. Psychological distress and mortality in systolic heart failure. Circ Heart Fail 2010;3:261-7.

[15] Damen N, Versteeg H, Boersma E, Serruys PW, van Geuns RJM, Denollet J, et al. Depression is independently associated with seven year mortality in patients treated with percutaneous coronary intervention: results from the RESEARCH registry. Int J Cardiol 2012 (Epub ahead of print).

[16] Volz A, Schmid JP, Zwahlen M, Kohls S, Saner H, Barth J. Predictors of readmission and health related quality of life in patients with chronic heart failure: a comparison of different psychosocial aspects. J Behav Med 2011;34:13-22.

[17] Grande G, Romppel M, Vesper JM, Schubmann R, Glaesmer H, Hermann-Lingen C. Type D personality and all-cause mortality in cardiac patients — data from a German cohort study. Psychosom Med 2011;73:548-56.

[18] Coyne JC, Jaarsma T, Luttik ML, van Sonderen E, van Veldhuisen DJ, Sanderman R. Lack of prognostic value of Type D personality for mortality in a large sample of heart failure patients. Psychosom Med 2011;73:557-62.

[19] Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosom Med 2004;66:411-21.

[20] Smith TW. Toward a more systematic, cumulative, and applicable science of personality and health: lessons from Type D personality. Psychosom Med 2011;73:528-32.

[21] Block J, Ozer JD. Two types of psychologists: remarks on the Mendelsohn, Weiss, and Feimer contribution. J Pers Soc Psychol 1982;42:1171-81.

[22] Humphreys LG, Fleishman A. Pseudo-orthogonal and other analysis of variance designs involving individual-differences variables. J Educ Psychol 1974;66:464-72.

[23] MacCallum RC, Zhang SB, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. Psychol Methods 2002;7:19-40.

[24] Maxwell SE, Delaney HD. Bivariate median splits and spurious statistical significance. Psychol Bull 1993;113:181-90.

[25] Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med 2006;25:127-41.

[26] Cohen J, Cohen P. Applied multiple regression/correlation analysis for the behavioral sciences. 2nd ed.Hillsdale, NJ: Erlbaum; 1983.

[27] Humphreys LG. Doing research the hard way — substituting analysis of variance for a problem in correlational analysis. J Educ Psychol 1978;70:873-6.

[28] Mendelsohn GA, Weiss DS, Feimer NR. Conceptual and empirical-analysis of the typological implications of patterns of socialization and femininity. J Pers Soc Psychol 1982;42:1157-70.

[29] Haslam N, Holland E, Kuppens P. Categories versus dimensions in personality and psychopathology: a quantitative review of taxometric research. Psychol Med 2012;42:903-20.

[30] Markon KE, Chmielewski M, Miller CJ. The reliability and validity of discrete and continuous measures of psychopathology: a quantitative review. Psychol Bull 2011;137:856-79.

[31] Ferguson E, Williams L, O'Connor RC, Howard S, Hughes BM, Johnston DW, et al. A taxometric analysis of Type-D personality. Psychosom Med 2009;71:981-6.

[32] Block J. A contrarian view of the five-factor approach to personality description. Psychol Bull 1995;117:187-215.

[33] Cortina JM. Interaction, nonlinearity, and multicollinearity: implications for multiple regression. J Manag 1993;19:915-22.

[34] Lubinski D, Humphreys LG. Assessing spurious "moderator effects": illustrated substantively with the hypothesize ("synergistic") relation between spatial and mathematical ability. Psychol Bull 1990;107:385-93.

[35] Grande G, Romppel M, Barth J. Association between Type D personality and prognosis in patients with cardiovascular diseases: a systematic review and meta-analysis. Ann Behav Med 2012;43:299-310.

[36] de Voogd JN, Sanderman R, Coyne JC. A meta-analysis of spurious associations between Type D personality and cardiovascular disease endpoints. Ann Behav Med 2012;44:136-7.

[37] Grande G, Romppel M, Barth J. Type D personality and heart disease: walking the line between enthusiasm and disbelief. Ann Behav Med 2012;44:138.

[38] Egger M, Schneider M, Davey Smith G. Meta-analysis — spurious precision? Meta-analysis of observational studies. BMJ 1998;316:140-4.

[39] Concato J, Peduzzi P, Holfold TR, Feinstein AR. Importance of events per independent variable in proportional hazards analysis. I. Background, goals, and general strategy. J Clin Epidemiol 1995;48:1495-501.

[40] Ketterer MW, Denollet J, Goldberg AD, McCullough PA, John S, Farha AJ, et al. The big mush: psychometric measures are confounded and non-independent in their association with age at initial diagnosis of ischaemic coronary heart disease. J Cardiovasc Risk 2002;9:41-8.

[41] Coyne JC. Self-reported distress: Analog or ersatz depression? Psychol Bull 1994;116:29-45.

[42] Ormel J, Rosmalen J, Farmer A. Neuroticism: a non-informative marker of vulnerability to psychopathology. Soc Psychiatry Psychiatr Epidemiol 2004;39:906-12.

[43] Suls J, Bunde J. Anger, anxiety, and depression as risk factors for cardiovascular disease: the problems and implications of overlapping affective dispositions. Psychol Bull 2005;131:260-300.

[44] Smith TW. Conceptualization, measurement, and analysis of negative affective risk factors. In: Steptoe A, editor. Handbook of behavioral medicine: methods and applications. New York, NY: Springer; 2010. p. 155-68.

[45] Grossardt BR, Bower JH, Geda YE, Colligan RC, Rocca WA. Pessimistic, anxious, and depressive personality traits predict all-cause mortality: the mayo clinic cohort study of personality and aging. Psychosom Med 2009;71:491-500.

[46] Meehl P. Why summaries of research on psychological theories are often uninterpretable. Psychol Rep 1990;56:199-244.

[47] MacLeod J, Davey Smith G, Heslop P, Metcalfe C, Carroll D, Hart C. Psychological stress and cardiovascular disease: empirical demonstration of bias in a prospective observational study of Scottish men. BMJ 2002;324:1247-51.

[48] Davey Smith G, Phillips AN, Neaton JD. Smoking as "independent" risk factor for suicide: illustration of an artifact from observational epidemiology? Lancet 1992;340:709-12.

[49] Macleod J, Davey Smith G. Psychosocial factors and public health: a suitable case for treatment? J Epidemiol Community Health 2003;57:565-70.

[50] Christenfeld NJS, Sloan RP, Carroll D, Greenland S. Risk factors, confounding, and the illusion of statistical control. Psychosom Med 2004;66:868-75.

[51] Flack VF, Chang PC. Frequency of selecting noise variables in subset regression-analysis — a simulation study. Am Stat 1987;41:84-6.

[52] Derksen S, Keselman HJ. Backward, forward and stepwise automated subset-selection algorithms — frequency of obtaining authentic and noise variables. Br J Math Stat Psychol 1992;45:265-82.

[53] Breslow N. Design and analysis of case–control studies. Annu Rev Public Health 1982;3:29-54.

[54] Albus C, Bjarnason-Wehrens B, Gysan DB, Herold G, Schneider CA, zu Eulenburg C, et al. The effects of multimodal intervention for the primary prevention of cardiovascular diseases on depression, anxiety, and Type-D pattern. Initial results of the randomized controlled PreFord trial. Herz 2012;37:59-62.

[55] Thombs BD, de Jonge P, Coyne JC, Whooley MA, Frasure-Smith N, Mitchell AJ, et al. Depression screening and patient outcomes in cardiovascular care a systematic review. JAMA 2008;300:2161-71.

[56] Lesperance F, Frasure-Smith N. The seduction of death. Psychosom Med 1999;61:18-20.

[57] Lesperance F, Frasure-Smith N. Negative emotions and coronary heart disease: getting to the heart of the matter. Lancet 1996;347:414-5.

[58] Park A. Are you a Type D personality? Your heart may be at risk. Time September 14, 2010 http://healthland.time.com/2010/09/14/a-new-risk-factor-for-heart-disease-type-d-personality/#ixzz1yutbt0Wl.

[59] Holroyd KA, Coyne J. Personality and health in the 1980s — psychosomatic-medicine revisited. J Pers 1987;55:359-75.

[60] Ioannidis JPA. Why most discovered true associations are inflated. Epidemiology 2008;19:640-8.

[61] Hemingway H, Riley RD, Altman DG. Ten steps towards improving prognosis research. BMJ 2009;339.

[62] Loder E, Groves T, Macauley D. Registration of observational studies. BMJ 2010;340.

[63] Begley CG, Ellis LM. Drug development: raise standards for preclinical cancer research. Nature 2012;483:531-3.

[64] Nosek BA, Spies J, Motyl M. Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. Perspect Psychol Sci (in press).

[65] Schroter S, Black N, Evans S, Godlee F, Osorio L, Smith R. What errors do peer reviewers detect, and does training improve their ability to detect them? J R Soc Med 2008;101:507-14.

[66] John LK, Loewenstein G, Prelec D. Measuring the prevalence of questionable research practices with incentives for truth-telling. Psychol Sci 2012;23:524-32.

[67] Bem DJ. Writing the empirical journal article. In: Darley JM, Zanna MP, Roediger HL, editors. The complete academic: a practical guide for the beginning social scientist. 2nd edition. Washington, DC: American Psychological Association; 2003.

[68] Caulfield T, Condit C. Science and the sources of hype. Public Health Genomics 2012;15:209-17.