The role of BDNF in depression : will the neurothrophin hypothesis sparkle on, long after the glitter of the firework is gone?

Molendijk, M.L.

CHAPTER 10

GENERAL DISCUSSION

The overarching purpose of this thesis was a better appreciation and a more refined model of (peripheral) neurotrophic functioning in depressive (and related) disorders. The empirical data of the prevailing thesis are presented in the previous eight chapters. The next chapter consists of an aggregation and a discussion of these data. This will be done first for the findings on the determinants of serum BDNF concentrations, to be followed by the clinical findings that are reported herein. Strengths and weaknesses will be discussed on the spot and additionally in a paragraph dedicated to this important subject. The discussion will continue with the possible implications of our findings and based on the current state of knowledge the main open questions will be stipulated. A summary in English, Dutch, German, and Spanish closes up this thesis.

Determinants of serum BDNF concentrations

Each empirical chapter in this thesis explored an area that either is directly relevant for our understanding of neurotrophic functioning in psychiatric illness, notably depression, or laid a basic to this end. **Chapters 2** and **3** are examples of the latter as here the determinants of serum BDNF concentrations were explored.

In line with a conceptualization of serum BDNF concentrations as being dependent on a myriad of factors, a long-list of variables was discovered that systematically are associated with serum BDNF concentrations. **Table 1** ↓ summarizes this list next to the main *null findings* that are reported in the **chapters 2** and **3**.

The Basic Determinants of serum BDNF concentrations

Chapter 2 addressed four categories of variables with regard to their association with serum BDNF concentrations. These categories were sampling-, socio-demographic-, lifestyle-, and disease variables.

Of the sampling variables (four were tested; time of blood withdrawal, the number of minutes that a sample was kept in a cool box before being processed, the duration of sample storage, and adherence to the fasting protocol) it was found, in multivariable analyses, that participants who were sampled later in the morning and those who did not adhere to the fasting protocol had lower serum BDNF concentrations. Longer sample storage of blood serum before BDNF determination took place also was associated with lower serum BDNF concentrations. Thus with time, even when stored at -85 C˚, the BDNF protein deteriorates. Herewith we replicate the findings of a Danish group (Trajkovska *et al.,* 2005) that addressed this topic for storage at -20 C˚.

The socio-demographic variables that were tested included gender, age, degree of urbanicity of living environment, and years of formal education. Of these, age was positively associated with serum BDNF concentrations. An age–gender interaction effect further specified this association and indicated that the increase in serum BDNF concentrations related to age was more strongly evident in females as compared to males. This finding was somewhat more complicated as the age related increase in serum BDNF concentrations in women seemed to end when the menopausal stage set in (around 50 years of age; Voorhuis *et al.,* 2011). A possible explanation for this will be provided in a latter part of this discussion. Living in a more urban area was associated with increased serum BDNF concentrations. With regard to the

lifestyle indicators (*i.e.,* smoking, alcohol consumption, the amount of physical activity, and body mass index) it was found that being an excessive drinker was associated with lower-, whereas smoking was associated with higher serum BDNF concentrations.

Table 1. Determinants of serum BDNF concentrations

Category	Variable	Relation to BDNF [1]
Sampling variables (chapter two, *N* = 1,168)	- time of morning blood draw	↓ when sampled later
	- sample in cool box before processing	no association
	- duration of sample storage	↓ when longer stored
	- fasting at times of blood draw	↓ when non-fasting
Socio-demographic variables (chapter two, *N* = 1,168)	- gender	no association
	- age (years)	↑ when higher age
	- level of urbanicity	↑ in urban areas
	- years of education	no association
Lifestyle indicators (chapter two, *N* = 1,168)	- physical activity	no association
	- alcohol consumption	↓ in excessive-drinkers [2]
	- smoking	↑ in smokers
	- body mass index	no association
Disease variables (chapter two, *N* = 1,168)	- metabolic syndrome	no association
	- chronic lung disease	no association
	- coronary artery disease	no association
Meteorological variables (chapter three, *N* = 2,851)	- season of blood sampling	↑ in late summer and early fall
	- month of blood sampling	↑ in September and October
	- ambient sunlight hours	↑ following sunny periods

[1] ↑, ↓: statistically significant higher or lower serum BDNF concentrations respectively
no association: no statistically significant association between serum BDNF concentrations and the variable that is indicated in the corresponding row
[2] excessive-drinking is defined as > 14 units per week

None of the disease related variables, including the presence of chronic non-specific lung- and coronary artery disease, were associated with serum BDNF concentrations. These null findings were not expected based on the literature (see for instance Golden *et al.,* 2010). A possible explanation for this may be the overall health status of the sample and, as a result, a lack of variation in the occurrence of these illnesses and thus low statistical power.

Seasonality in serum BDNF concentrations

Chapter 3 assessed seasonality, a broad concept that refers to biochemical or behavioral alternations as response to variability in the length of day (Walton *et al.,* 2011), in serum BDNF concentrations. A drive to study this topic was that in rodents seasonality has been observed in neuronal plasticity, a process that is regulated by BDNF (Workman *et al*., 2009). A further motivation was that molecular events, upstream of BDNF have been shown to undergo seasonal rhythmicity (Lambert *et al.,* 2002). Finally, depression presumably is related to serum BDNF concentrations (Sen *et al.,* 2008) and this illness occurs to some extent according to a seasonal pattern (Lewy *et al.,* 2006). So, seasonality in serum BDNF concentrations

was expected. And indeed, particular strong evidence for our expectation was found: serum BDNF concentrations increased in the spring-summer period and decreased in the autumn-winter period. This effect was independent of potential confounders such as having a DSM-IV depression diagnosis. Importantly, the observed effect-size estimates for monthly differences in serum BDNF concentrations were substantial (up to a Cohen's *d* of 0.60). Explorative analyses further showed that the number of sunshine hours (a major trigger to entrain seasonality; Walton *et al.,* 2011) in the weeks prior to blood withdrawal positively correlated with serum BDNF concentrations and this partly explained the observed monthly variation. It was also found, and this may not come as a surprise now, that the length of day correlated in a similar manner with serum BDNF concentrations as the number of ambient sunshine hours did.

The findings on the determinants of serum BDNF concentrations have significant implications, as will be discussed in the part that follows.

Determinants of serum BDNF concentrations: implications

Knowledge on the determinants is of importance. Smith and Ebrahim (2002, page 1438) wrote that *we live in an associational world* where people who differ in some regard from others, often differ systematically in other regards as well (*e.g.,* persons who are depressed are more likely to smoke [Kendler *et al.,* 1993]). So, characteristics cluster together and as such they may have shared relationships with certain outcomes. Confounding is said to occur then when one element of a cluster is associated with a given outcome, whereas this relation is due to another element of the cluster. Confounding is the most likely cause of spurious associations (Smith and Ebrahim, 2002) and as such a stand in the way of (research) progress. Gaining detailed insight into determinants or the confounding structure of certain traits, biological alternations or behavior is a means to avoid this because it can provide a scaffold for the exploration of independent associations. The findings in the first two chapters of this thesis provide such a scaffold.

Determinants of serum BDNF concentrations: methodological implications

First, given that the time of the day of blood draw and non-fasting protocol are associated with serum BDNF concentrations, a stringent sample protocol is warranted to obtain valid results. In this protocol a narrow time range should be defined in which blood sampling should take place (*e.g.,* between 07:00 and 07:30 a.m.) and it should be specified and controlled for that participants are sober at the time of blood draw.

Next, since the duration of sample storage impacts on serum BDNF concentrations, when studying differences among diagnostic groups, one needs to make sure that the groups that are compared do not differ in a systematic manner. Let me provide an (oversimplified) example for why this is/becomes (increasingly) important. The US army stores blood from all its soldiers since the start of the first Gulf war in 1991 (*Nature News*, July 2013) to answer relevant questions with. Say, one wishes to learn whether veterans who were sent to Iraq and who did and did not develop post-traumatic stress disorder differ from non-soldier controls with and without PTSD with regard to serum BDNF concentrations. Data on the soldier group comes from the US army database whereas those on the control group need to be gathered after the research question is formulated (let's say June 2014). Given that BDNF levels decrease about 1/10 of a standard deviation each year as a function of storage time (**chapter 2**) the amount of BDNF in the serum of the soldiers (sampled in 1991) would at least be 2 standard deviations lower compared to those of the non-soldier controls (sampled in 2014). Concluding that soldiers who develop post-traumatic stress disorder have lower serum BDNF concentrations as compared to healthy controls with PTSD obviously wouldn't be valid as the between-group difference is due to storage duration.

The discovery of seasonality in serum BDNF concentrations (**chapter 3**) is important for methodological reasons alike. One, it is crucially important when interpreting the results from longitudinal studies. In fact, trials that span months and that have serum BDNF concentrations as an outcome, may be of little use unless detailed knowledge on seasonal effects is taken into account (in the protocol and/or the statistically). Likewise, there is the need to sample groups (*e.g.,* depressed versus healthy controls) equally over the year in order to gain credibility in research findings. Given the seasonal patterns in the occurrence of mood disorders (Lewy *et al.,* 2006) this is quite difficult. I will illustrate this with an example that applies to one of the main findings that is presented in this thesis; the lower serum BDNF concentrations in depressed persons as compared to healthy controls (**chapter 4**, to be discussed in a next part). In the NESDA sample, persons who are depressed are more likely to be sampled in the winter (42 percent of the participants) as compared to the summer (36 percent of the participants). This 6 percent difference may not seem that large but it is statistically significant (*P*-value for the difference = .001) and given the large sample-size of the NESDA a difference of a few percent involves several dozens of persons. Just as the prevalence of depression, serum BDNF concentrations differ as a function of season, with higher concentrations in persons who are sampled in the summer as compared those who are sampled in the winter (d = 0.47, P < .001, see **chapter 3**). Together this already suggests that controlling for season of sampling may make a difference. This indeed is so. Our data shows that the differences in serum BDNF concentrations between depressed patients and healthy control subjects is statistical significant (P = .007) and has an effect size of d = 0.19. When this analysis is rerun controlled for seasonality, the direction of the effect and its statistical significance hold, yet the latter shifts upwards to P = .04 and importantly, the strength of the association is attenuated by about 40 percent to d = 0.11. This is illustrative for the importance of accounting for determinants. Please note that large sample size is more robust with regard to deleterious effects of confounding as compared to studies that conclude on the basis of a smaller sample (Lenth 2001). Therefore the effect-size estimates derived from small-scale studies may be hampered by confounding to a larger extent and thus the results that these yield may not only be less reliable but also less valid.

Concluding from the above, a stringent sample protocol and/or statistical control for a range of relevant variables seem warranted in BDNF related research in order to come to valid results. I gladly noticed that, based on our findings, several authors have picked in their studies on serum BDNF concentrations (*e.g.,* Ball *et al.,* 2013). What the above-presented findings obviously do not bring is insight into how exactly serum BDNF concentrations vary as a function of different levels of *determinant exposure*. this however was not the intention of this thesis and the epidemiological nature of the data that were used gave little room for studying this. Still, some of the findings herein do hint to mechanisms that govern BDNF expression. Given that I consider such hints important in generating future hypotheses, I do not want to simply jump over them. Two illustrative examples therefore are discussed below: (**I**) the interaction effect between age and menopausal stage and serum BDNF concentrations (**chapter 2**) and (**II**) the delayed positive linear relationship between the number of ambient sunlight hours and serum BDNF concentrations (**chapter 3**). Please note that we did not formally test the hypotheses that are brought forward in the following section.

Gender specific associations: estrogen and BDNF

Chapter 2 reports a larger to age related, increase in serum BDNF concentrations in women as compared to men. Given the large number of participants that formed the base for this result (757 women, 408 men), this is not likely to be a chance finding. In addition, we could convincingly exclude that behavioral- or illness characteristics were the source of this association. What could have caused it? Hormonal differences between the two genders seem a viable explanation. In accord with this was that the age related rise in

serum BDNF concentrations in women occurred until the menopausal stage (~ 50 years of age; Voorhuis *et al.,* 2011) and not thereafter. In fact, in female participants it was observed that from ~ 50 years on, age was not associated with serum BDNF concentrations. So, serum BDNF concentrations peak at the climacteric age in women. This is interesting because (**I**) there are studies that have established positive correlations among peripheral estrogen and BDNF concentrations (*e.g.,* Monteleone *et al*., 2007) and (**II**) estrogen expression drops sharply in females in the menopausal stage (Genazzani *et al.,* 1999). Based on this, we hypothesized that the association between age and serum BDNF concentrations could be dependent on menopausal stage and by extension, maybe, on estrogen expression. To probe this further, women in the age range 48 to 52 years were selected and the interaction between age and menopausal status on serum BDNF concentrations was modeled. This analysis showed that in premenopausal women BDNF concentrations increased as a function of age whereas in post-menopausal such a relation was not observed. This corroborates, although not proves, the idea that in women serum BDNF concentrations are under the influence of estrogen expression. This may have clinical implications in that the transition into the menopause is associated with increased odds on depression (Judd *et al.,* 2012) and a relatively large drop in cognitive performance (Farrag *et al.,* 2002), which, according to the neurotrophin hypothesis, are both under the influence of BDNF (Duman *et al.,* 1997). Interestingly, and in line with this idea, is that in **chapter 6** we find that anxious women have lower serum BDNF concentrations as compared to healthy women, whereas in male patients with an anxiety disorder this effect was not observed. This gender specific association could, in theory, also be explained by estrogen because the expression of this hormone is low in anxious women (Walf and Frye, 2006). Some pre-clinical studies suggest the potential importance of this on the level of the central nervous system by showing that estrogen - BDNF interactions are associated with dendritic growth and synaptogenesis in the cerebellum (Haraguchi *et al.,* 2012). So, estrogen and BDNF interactions may be of importance in understanding age and gender related changes in behavior (*e.g.,* depression) and abilities (*e.g.,* cognitive performance). Of course, a manifold of changes occur during the menopausal transition that were not controlled for so alternative explanations loom. Besides, estrogen does not explain some other observations of our studies (*e.g.,* the increase in serum BDNF concentrations as a function of age in males).

Let the sun shine bright: serotonin and BDNF

Our study described in **chapter 3** also yielded some results that hint to mechanisms that govern BDNF expression. Here, I point to the positive relationship between the number of ambient sunlight hours and serum BDNF concentrations. Interestingly, this effect was observed with a delay; where the number of ambient sunlight hours in the weeks prior to blood draw (for up to 7 to 8 weeks before this event) correlated positively with serum BDNF concentrations, the number of sunlight hours in the week of the blood draw itself were not. Importantly, also for this finding we could exclude a range of confounding factors such as the time of the day of blood draw and the study that yielded this finding was well powered.

A possible explanation for this effect can be gained through a comparison with an observation from the treatment setting. Preclinical and clinical studies have shown that antidepressants upregulate the transcriptional activity of BDNF (Duman and Monteggia, 2006) but only after long-term administration (*e.g.,* ≥ 21 days; Conti *et al*., 2002). The gist on why the increase in BDNF expression occurs with a lag time is that antidepressants first increase the availability of monoamines, notably serotonin, to set in CREB activity (Castren *et al.,* 2007). CREB, a general transcription factor (Guilloux *et al*., 2012), binds to the promoter region of the BDNF gene and this positively regulates transcription (Impey *et al*., 2004). Therefore it seems likely that (serotonergic) antidepressants act on CREB through an increase in the availability of serotonin (Impey *et al.,* 2004). Just like treatment with pharmacological antidepressants, long day conditions give rise

to a higher expression of serotonin (Lambert *et al*., 2002). This increase is very likely to occur through a direct effect of light on the expression of L-Tryptophan. L-Tryptophan is known to increase 5-Hydroxy-tryptophan and serotonin expression, leading in turn a higher level of CREB activity and BDNF expression (Castren *et al.,* 2007). Therefore it seems likely that also the delayed increase in serum BDNF concentrations following relatively sunny periods is entrained by increased CREB activation that is induced by a larger availability of biological active serotonin and its precursors L-tryptophan (Cappielo *et al.,* 1996) and 5-hydroxy-tryptophan (Wehr *et al.,* 2001).

An final note on the finding of seasonality in serum BDNF concerns that, although here described as a confounder, it in theory also could be regarded as a mediator linking season and depression. This mediator role could not be excluded in our studies, although it may not seem that likely given the lag-time with which the change in BDNF occurred and the absence of a depression diagnosis times season interaction effect on serum BDNF concentrations.

The determinants of serum BDNF concentrations – recapitulating

In my view, the data that the first two chapters of this thesis bring improve the base to understand inter-individual differences in serum BDNF concentrations. Besides, the acquired knowledge will facilitate ongoing research into neurotrophic functioning in depression (and related illnesses). It allowed us to test, largely independent, some predictions from the neurotrophin hypothesis, which is the topic of the next part of this discussion.

The neurotrophin hypothesis of depression and our work

The rationale for the neurotrophin hypothesis of depression is straightforward: BDNF expression, shaped by genetic and environmental influences, determines neuronal faith and viability and subsequently behavior, including depression (Duman *et al.,* 1997). The two basal predictions from this hypothesis are that depression results from a stress-induced decrease in BDNF expression and that antidepressants are efficacious because they normalize this (Duman and Monteggia, 2006). The overarching purpose of this thesis was to evaluate the validity of these predictions using peripheral BDNF parameters and a genetic variant that is presumed to be associated with neurotrophic functioning. In the section below I will discuss the studies that were performed to this end.

Serum BDNF concentrations in depressive illness

Chapter 4 and **5** advanced the understanding of the associations between serum BDNF concentrations and the illness major depression. **Chapter 4**, a single study, reports in accord with previous findings (Sen *et al*., 2008) and the neurotrophin hypothesis (Duman *et al*., 1997) that serum BDNF concentrations are low in depressed patients as compared to healthy controls (d = 0.19). Importantly, this study also shows that serum BDNF concentrations are low in depressed patients as compared to persons who are in full remission and that serum BDNF concentrations of this latter group are comparable to those of controls. Thus, low serum BDNF concentrations are a state characteristic of depression; an abnormality that is evident during depression and that normalizes during remission.

Chapter 5, a meta-analysis on the same subject, establishes the robustness of this association as it shows, based on 2,384 depressed patients and 2,982 healthy controls that serum BDNF concentrations are low in the depressed state (d = -0.47, 95% *CI* = -0.64 — -0.27). Herewith the neurotrophin hypothesis is corroborated in its prediction that serum BDNF concentrations are abnormally low in depressed patients. What may be the cause of this between-group difference? The axiom that has been brought forward by the neurotrophin hypothesis is trauma- or stress exposure.

Serum BDNF concentrations and trauma/stress exposure

The axiom that for over a decade has been brought forward in explaining depression related alternations in BDNF expression is early life trauma- (such as childhood abuse) or stress exposure in adulthood (Duman *et al.,* 1997; Duman and Monteggia, 2006). Some preclinical and clinical evidence exists for this idea (see for instance Groves 2007). In contrast, the well-powered and well-controlled studies in this thesis show that trauma exposure is not associated with BDNF concentrations (**chapter 7**). In fact, with regard to stress exposure we only could show a negative correlation between recent stress exposure (*e.g.,* a divorce) and serum BDNF concentrations. Given that this association only explained ~ 1 percent of the variance in serum BDNF concentrations it probably does not constitute a sufficient explanation for alternations in neurotrophic functioning in stress related illnesses. The relationship between trauma exposure and BDNF, if it is truly there, probably is more complex and for instance dependent on the presence of a moderator (see **chapter 7** in which a val^{66}met - trauma interaction effect on serum BDNF concentrations was established). This will be discussed in a later part of this discussion. Nothwitstanding a lack of knowledge on moderators, I wish to state that the findings regarding trauma- and stress exposure largely, as they appear to be now (*i.e.,* largely negative) are not in line with the neurotrophin hypothesis. So, maybe other explanations need to be sought for the lower BDNF concentrations in stress related illnesses such as depression. Some findings in this thesis suggest on such mechanisms, other then stress. These putative mechanisms, alongside some some findings that show a lack of fit with the predictions of the neurotrophin hypothesis, will be the topic of the section below.

Serum BDNF concentrations and the early remission phase: reverse causation?

In **chapter 4** we found that patients who were in the early remission phase of their depressive episode, and thus largely free of symptoms, had serum BDNF concentrations that were lower as compared to those of currently depressed patients. Thus, serum BDNF concentrations remain low, or even become somewhat lower, after clinical improvement has set in. Explanations for this could not be elucidated in this thesis and longitudinal designs with frequent samplings need to be performed to understand this issue. Albeit the lack of a clear explanation, the finding of low levels of BDNF in the early remission phase is not in line with the prediction of the neurotrophin hypothesis that low neurotrophic support endangers a person to become depressed (Duman and Monteggia, 2006). In fact, it seems to plead for *reverse causation* in that the lower serum BDNF concentrations do not endanger a person to become depressed but rather are a consequence of being depressed. Why would serum BDNF concentrations be particularly low in the early remission phase of depression? Two hypotheses will be formulated here that potentially can explain it: (**I**) to depression related changes in body-weight and (**II**) to depression related changes in levels of oxidative stress. These hypotheses will be discussed below.

In **chapter 4** we find that BMI is positively associated with serum BDNF concentrations in depressed patients (note that such a correlation was not observed in healthy control subjects, see **chapter 2**). Although this finding was unsought, it parallels the results of some previous studies (Nakazato *et al*., 2003; Monteleone *et al*., 2005) and they give ground to an interesting hypothesis. As weight loss is a prime behavioral abnormality of depression (APA 1994) and often a residual symptom in early remission (Paykel 1985; Paykel *et al*., 1995) it could be that alternations in serum BDNF concentrations are mediated by (transient) changes in eating behavior during, or in the aftermath of a depressive episode. The mechanism would then simply be that a decrease in intake of the building blocks for the protein BDNF could lower the expression of it or in this case particular, a higher metabolism of BDNF and consequently lower BDNF concentrations. Likewise, weight gain is a documented side effect of antidepressant treatment (Kachur *et*

al., 2005). And this, or better the absence of weight loss could potentially explain the absence of a relative fall of serum BDNF concentrations in depressed patients who are treated with an antidepressant (the associations among antidepressants and serum BDNF concentrations will be discussed in a latter part of this discussion).

Another explanation for the abnormally low serum BDNF concentrations in the early remission- and the active depression phase is mediation by oxidative stress. Depressive disorders are accompanied by a decreased antioxidant status (Maes *et al.,* 2011). The antioxidant status of a person refers to the capacity to protect against reactive oxygen species. An imbalance of the oxidative status generates toxic reactive oxygen species and this causes damage to membrane lipids, to DNA and consequently disturbs the functioning and stability of proteins (Sarandol *et al.,* 2007). So, oxidative stress causes oxidative imbalance with accompanying protein damage and also BDNF functioning may be negatively affected by it. Alternations in oxidative stress homeostasis set in during the depressed state (probably due to behavioral alternations such as a changed eating pattern and less physical activity). Therefore also the lowering of BDNF functionning, if it truly is affected by oxidative stress, may only set in during the depressed state (critically, not prior to the depressed state as the neurotrophin hypothesis suggests). Given that the disturbed oxidative stress homeostasis may linger on into early remission (Barnham *et al.,* 2004), BDNF levels consequently may remain low in this phase. There is one study in human subjects that confirmed the idea on the role of oxidative stress on neurotrophic functioning to some extent. This study by Kapczinski *et al.* (2008) showed a negative correlation (r = -0.58) between serum thiobarbituric acid reactive substances and BDNF concentrations in bipolar patients (in whom serum BDNF concentrations in general are low; Fernandes *et al.,* 2011). This is suggestive for the notion that alterations in oxidative stress homeostasis may be mechanistically associated with alternations in the expression/metabolism of BDNF. There are some preclinical studies that support this idea. Already in 1996, Kirschner and her colleageaus showed, *in vivo*, that neuronal damage and decreased BDNF expression can be induced by chemical hypoxia. This, however, could be attenuated by BDNF administration. Interestingly, some authors have suggested that antidepressants may affect oxidative stress homeostasis in a positive manner (Khanzode *et al.,* 2003). Therefore, normalization of oxidative status may complementary explain the absence of abnormally low serum BDNF concentrations in antidepressant treated depressed persons. Finally, is notable that several studies show that oxidative stress is associated with processes that typically are governed by BDNF, such as neuronal functioning (Barnham *et al.,* 2004).

Whatever the mechanism that is involved in the lowering of serum BDNF concentrations in the depressed and the early remission phase, **chapter 4** suggests that the effect-size on these differences are small in absolute sense (*i.e.,* a standardized mean difference of ~ 0.2) and also when compared to those reported by earlier small scale studies and meta-analyses (see for instance Sen *et al.,* 2008). Small effect-sizes indeed were confirmed in the meta-analysis, *the way to converge to the true effect-size*, in **chapter 5**. The findings of this chapter and the implications that they may have are discussed below.

Serum BDNF concentrations and depression diagnosis – small effect-size estimates

Chapter 5, a large-scale meta-analysis, shows lower serum BDNF concentrations in untreated depressed patients as compared healthy controls. This finding is not new (see above). The novelty of the work instead is that it highlights a large amount of between-study heterogeneity in outcomes. Importantly, none of the theoretically relevant variables that we tested (*e.g.,* gender distribution of the sample) was associated with the between-study heterogeneity. Obviously, it may have come from between-sample characteristics, such as heterogeneity in clinical characteristics of patient samples. However, for this idea, meta-regression analyses could find no evidence whatsoever. In contrast, these analyses showed an artificial base for the

heterogeneity in outcomes. First, a large part of the studies that were included was underpowered. Given that a low level of power increases the false versus true positive ratio (Sterne and Smith, 2001), some overly positive findings may have been among the studies that we included, causing heterogeneity. Second, we found that sample size and year of publication were significant predictors of between-study heterogeneity, with larger and more recently published studies reporting smaller between-group differences. This points to publication bias – which is a threat to the validity of the literature and besides a cause of hererogeneity. Analyses that accounted for publication bias yielded an attenuated effect-size estimate that was about half as large (*i.e., d* = 0.47) as the one reported in previous meta-analysis that was based on 8 times less data (Sen *et al.,* 2008). Of course, small-effect sizes do not attest the validity of a hypothesis. In fact, the effect-size estimate remained statistically significant and thus corroborates it. What this finding does attest however is that the data, for reasons of publication bias and overestimations of effect-sizes, should be critically interpreted. A consequence of this is that the evidence for the neurotrophin hypothesis is slimmer as was initially thought. This is in line with more recent work showing that, using longitudinal designs, serum BDNF concentrations are not, or only marginally, predictive for depression related psychopathology (*e.g.,* Vinberg *et al.,* 2013; Bus *et al.,* submitted). Furthermore, the findings in this thesis on the lack of an association between serum BDNF concentrations and the clinical features of depression and all major types of antidepressants use (**chapters 4** and **5**) also suggest that the evidence for the neurotrophin hypothesis is not so strong as was initially thought. This will be discussed in below.

Serum BDNF concentrations and the clinical features of depression

The putative association between serum BDNF concentrations and the symptom severity of depression has been brought forward as a pillar of the neurotrophin hypothesis (*e.g.,* Karege et al., 2002). This makes sense because when you predict that serum BDNF concentrations play a role in the pathophysiology of depression as the neurotrophin hypothesis does (Duman and Monteggia, 2006) you may expect that BDNF levels are are particularly low in more severely depressed patients. The findings reported in **chapter 4** and **5** however are not in line with this expectation. The single study reported in **chapter 4** could not replicate the association between depression symptom severity and serum BDNF concentrations in unmedicated depressed persons (r = 0.03, P = 0.23, N = 541) neither could the meta-analysis in **chapter 5**. The meta-analysis added a layer of certainty to this, as it showed that the studies that did find evidence for this association are outliers and likely false positives. So, the dose-response association between serum BDNF concentrations and the symptom severity of depression probably does not exist. This further is exemplified by the finding in **chapter 4** that the early remission phase is accompanied by a much lower symptom severity of depression (mean depressive symptom severity scores were 22.4 ± 11.4 versus 32.4 ± 12.1 (P < .000001) in currently depressed patients) alongside lower BDNF concentrations as compared to the current depressed state.

Regarding the other clinical characteristics of depression, **chapter 4** also could not confirm the findings that having a recurrent compared with a first episode of depression (Lee *et al*., 2007) and the presence of suicide ideation (Deveci *et al*., 2007; Kim *et al*., 2007) are accompanied by lower concentrations of BDNF. Age at onset of depression, the presence of comorbid anxiety and the chronicity of depression also were shown to be unrelated to serum BDNF concentrations (see **chapter 4**). Because most of the studies in the literature did not report on these variables there was no opportunity to confirm the null-findings regarding these clinical characteristics through meta-analysis. Notwithstanding this, together these findings, given the samples-sizes that were used to come to them, give confidence in excluding the clinical features of depression as potential correlates of serum BDNF concentrations. This is an important conclusion, as it

hints that other factors than specifically to depression related phenomenon, such as weight gain and loss, may be at play in the relative fall of BDNF concentrations during/around a depressive episode.

Serum BDNF concentrations and antidepressant treatment

A core prediction of the neurotrophin hypothesis is that antidepressants are clinically efficacious because they normalize neurotrophic functioning and with this aberrant brain functioning (Duman and Monteggia, 2006). In the **chapters 4** and **5** we find some evidence for this prediction.

Within the context of a large-scale single study (**chapter 4**) we found evidence that serum BDNF concentrations normalize in the course of antidepressant treatment. This finding largely is in accord with previous findings (Brunoni *et al*., 2008). It should be noted though that the effect-size on this association is considerably smaller (*i.e., d* = 0.23) as compared to the effect-sizes that previous studies report on. Still, in **chapter 5** the reliability of this finding was confirmed through meta-analysis, albeit this analysis also yielded a considerably smaller effect-size estimate (*i.e., d* = 0.34) as compared to those in previous studies (*i.e., d* = 0.80; Sen *et al.,* 2008). Besides, we were able to expand previous findings by showing that the use of an antidepressant is associated with increased serum BDNF during a depressive episode but not during remission. This suggests that antidepressant-induced increases in BDNF occur in a disease state when BDNF functioning might be defective and not in full remission when BDNF functioning is normalized (see **chapter 4**).

Interesting was that the increase in serum BDNF concentrations appeared to be specifically associated with the use of SSRIs and St John's wort and not with the use of SNRIs, TCAs or NaSSAs (**chapter 4**). Although not directly confirmed, this finding might be explained by increased availability of extra-synaptic concentrations of serotonin, as it is known that serotonin stimulates the expression of BDNF (Martinowich and Lu, 2008). In line with this, we found the highest BDNF concentrations in patients who were treated with agents that generally lead to an increase in the availability of serotonin; SSRIs and St John's wort (Gaster and Holroyd, 2000). Furthermore, the lowest concentrations of BDNF were found in patients who were treated with NaSSAs that are known to have little or no impact on the availability of serotonin (Antilla and Leinonen, 2001). Noteworthy is that this observation, and its putative explanation in terms of the availability of serotonin fits very well with the findings on seasonality in serum BDNF concentratons (see **chapter 3**). The antidepressant specific effect on serum BDNF concentrations however could not be replicated in the meta-analysis reported in **chapter 5**. This could be due to a lack of statistical power as in this analysis it had to be assessed through a meta-regression that used the number of included study as data-points and this number was only 28.

Notwithstanding the findings that seem to confirm the neurotrophin hypothesis, the results that were described in **chapter 4** and **5** do not all recapitulate it. First, the antidepressant-specific effect on serum BDNF concentrations (already mentioned above) seems at odds with the specific prediction of the neurotrophin hypothesis that increases in BDNF concentrations are a key mediator for an antidepressant response to occur (Duman and Monteggia, 2006). According to this prediction, one might expect that antidepressants that are known to be about equally efficacious in the treatment of the symptoms of depression (Berton and Nestler, 2006) would have similar effects on serum BDNF concentrations. Clearly, this is not the case.

A second finding that seems hard to reconcile with the neurotrophin hypothesis is that the group of depressed persons who used antidepressants (for prolonged period and on a frequent base) had the highest BDNF concentrations, but also the highest symptom severity of depression (**chapter 4**). This suggests that increases in peripheral BDNF concentrations do not parallel clinical effectiveness, or at least have no direct effects on depression characteristics such as its severity. Thus, the temporal dynamics as

predicted from the neurotrophin hypothesis do not seem to be correct. Furthermore, epiphenomena for the finding that antidepressants are associated with serum BDNF concentrations cannot be excluded. Interestingly, among these epiphenoma are those that also could explain differences among untreated depressed persons and healthy control subjects (*i.e.,* to depression and treatment related weight gain and loss and oxidative stress levels). Finally, **chapter 5** reports, along a similar line as the findings regarding differences among diagnostic groups, that a large amount of unexplained between-study heterogeneity in outcomes and publication bias is evident in the literature on the association between antidepressants and serum BDNF concentrations. This also indicates that the effect of antidepressants on serum BDNF concentrations is not that large as they initially were thought to be.

Relevant to the above-presented discussion on lower BDNF concentrations in persons diagnosed with depression are the findings reported in **chapter 6**. In this chapter serum BDNF concentrations were evaluated in persons diagnosed with an anxiety disorder. Given that the anxiety disorders mimic depressions to a great extent (David et al., 2009), abnormalities in serum BDNF concentrations were expected. This putative association is the topic of the section that follows.

Serum BDNF concentrations and anxiety

Based on animal models (*e.g.,* Monteggia *et al*., 2007), some small-scale human studies (Strohle *et al*., 2010), and the neurotrophin hypothesis (Duman and Monteggia, 2006) there is a strong a priori reason to expect that serum BDNF concentrations are low in persons who suffer an anxiety disorder. However, robust evidence for this belief is absent. **Chapter 6** filled this gap and explicitely tested the hypothesis that serum BDNF concentrations are low in patients with an anxiety disorder as compared to healthy controls. The results of this study, controlled for a range of demographical and behavioral variables and derived from a sufficiently powered design however did not confirm this hypothesis as overall no differences in serum BDNF concentrations between patients with an anxiety disorder and healthy controls were found. So, it seems unlikely that BDNF is involved in the pathophysiology of anxiety disorders *per se*.

Given that the preclinical literature gives ground to test for gender differences in the relation between anxiety and BDNF (Govindarajan *et al*., 2006), we assessed gender differences as well. Analyses stratified by gender revealed that female patients had somewhat lower BDNF concentrations relative to female controls (d = 0.19), whereas BDNF concentrations were similar among male patients and male controls. This gender specific finding could point in the direction that BDNF is related to the pathophysiology of anxiety in female but not in male patients. Remarkable is that this finding compares well with some studies from the depression literature, which have shown lower concentrations of BDNF in female depressed patients as compared to male depressed patients (Karege *et al*., 2002a; Huang *et al*., 2008; and **chapter 4**).

The origins of this gender specific finding are unknown. Here, and also in depressed subjects (see **chapter 4**) we found that the differences in serum BDNF concentrations between female and male patients were not driven by demographical (*e.g.,* age), behavioral (*e.g.,* smoking), or clinical (*e.g.,* severity) variables. In **chapter 6** it further was shown that the difference between female and male patients could not be attributed to a specific subtype of anxiety. A general deduction from this, and from the finding that serum BDNF concentrations are similar among female and male controls, is that the origins of it may lie in a female specific associate of anxiety. One interesting candidate that might serve as an explanation is the expression of the ovarian hormone estrogen, which in women with an anxiety disorde typically is low (Seeman, 1997; Almeida *et al*., 2005; Walf and Frye, 2006). This is of relevance here because estrogen is a signaling molecule that triggers the expression of BDNF (Scharfman and MacLusky, 2004; Begliuomini *et al*., 2007). This explanation relates to the observation in **chapter 2** that in women serum BDNF concentrations seem to peak at the climateric age (which already is discussed in an earlier part of this discussion).

Alternatively, and also not unlikely, the lower BDNF concentrations in female patients may be a female specific artifact of being anxious that is (causally) unrelated to the disease state itself. Therefore, and because of the findings that males did not show anxiety related lower serum BDNF concentrations, I conclude that it is unlikely that BDNF is involved in the pathophysiology of anxiety disorders *per se*. This conclusion is in contrast to the neurotrophin hypothesis (Duman and Monteggia, 2006) that patients with to depression related conditions exhibit abnormally low neurotrophic support. A final finding from **chapter 6** that deserves to be noted is that a history of depression seems unrelated to serum BDNF concentrations in patients with a current anxiety disorder, which corroborates our finding that low serum BDNF concentrations are a state characteristic of depression (see **chapter 4**).

The gender specific association and the low effect-size estimates that are reported in this thesis bring me to the topic of clinical utility. Studies into neurobiological abnormalities in psychiatric illness, BDNF in the prevailing thesis, may serve two functions: (**I**) they may help to parse out the pathophysiology of a certain illness condition and (**II**) they may add in the classification of a diagnostic condition or in the prediction of how successful a given treatment will be. The second function*, clinical utility*, will be discussed in some detail in the section that follows.

Clinical utitlity: diagnostic and treatment biomarkers

I will start this section with a short introduction on the concepts biomarker, moderation and mediation.

As stipulated in an earlier part, depressive disorders nowadays are diagnosed based on subjective assessments. This comes with disadvantages, as they may be inaccurate and/or colored by the state a patient is in. Therefore it is generally believed that an objective biological marker could improve the diagnostic process (*i.e.,* a *diagnostic biomarker*; Schmidt *et al.,* 2011). In addition, a biomarker could help to reduce heterogeneity by classification in a finite number of illness subtypes, which could have as advantage that patients can be assigned to treatment options that best fit their needs (Schmidt *et al.,* 2011). This latter distinction can be regarded as a treatment *moderator*; a differentiation that is used to decide for whom a certain treatment will work: depression type A → antidepressant X → depression alleviation *versus* depression type B → antidepressant Y → depression alleviation. Several biological markers have been studied in relation to depression (*e.g.,* cytokines, metabolic markers; see Kapur *et al.,* 2012), yet so far without success in that none of them have led to true clinical gain. A well-known example of this is the dexamethasone suppression test, which in the 1970/80s had initial promise in predicting relapse into depression. After extensive and sufficiently powered studies, this test however appeared to have a rather low sensitivity (~ 40-50%) and specificity (~ 70%)(APA taskforce on laboratory tests in psychiatry, 1987) and therefore could not be translated in clinical utility.

The concept of treatment mediator can be described along a similar line. A mediator however is different from a moderator in that it describes a mechanism by which a treatment may work (*i.e.,* a treatment mediator: depression → treatment → increase in biomarker X → depression alleviation). Information on treatment mediators also would be of clinical help in that objective assessment early in the course of treatment may be used as a marker for (early or future) treatment success and hence may improve drug efficacy (Schmidt *et al.,* 2011) and the understanding of the mechanisms that underlie antidepressant action.

BDNF as biomarker

Could serum BDNF concentrations possibly serve a biomarker function? Around the time that I started my thesis (2008) the answer on this question was *yes*. Sen *et al.* (page 527) for instance, concluded their 2008 meta-analyses on 748 subjects of whom 366 were depressed that serum BDNF concentrations *may have*

use as a biomarker for major depressive disorder or antidepressant efficacy. The effect-sizes on which this conclusion was based (*i.e.,* Cohen's *d* ~ 0.80) made their conclusion seemingly valid. However, and looking fairly at an effect size of 0.80, what does it tell us and can it be applied to distinguish between-groups in a diagnostic setting? Yes, it can, but only to a certain extent (i*.e.,* probability that an individual is correctly classified based on it (only) is 0.66 [Coe 2002]).

We report in **chapter 4** and **5** effects sizes on between-group differences that are considerably smaller (*e.g.,* for the difference between untreated depressed persons and healthy controls the best estimate is *d* = -0.47 (95% *CI* = -0.64 — -0.27). How well can a person be assigned to a group based on such an estimate? Not so well. In fact, the probability of correct classification is .59. Please note that an effect size of 0.00 would yield a .50 correct classification. Given that the relevance of a diagnostic biomarker depends on the magnitude of an effect-size (and not on statistical significance), I conclude that serum BDNF concentrations are lof little, if any, clinical use.

What is more is that lower BDNF concentrations have been reported in persons diagnosed with schizophrenia, with bipolar disorder, with eating disorders, etcetera. In **Table 2 ↓** (next page) the most reliable evidence (*i.e.,* derived from the largest single study or from meta-analysis) for alternations in serum BDNF concentrations in some psychiatric and neurological illnesses is presented. From this table it becomes clear that low serum BDNF concentrations are not specific enough to differentiate among diagnoses. So arguably, these values are very little informative in the clinical setting. Another reason why the issue of BDNF as a diagnostic biomarker may need a nuanced approach is that basically all the findings are acquired from between-subjects designs and therefore the data apply to groups and not to individuals. Furthermore, for the clinical use of BDNF parameters the detailed knowledge on the myriad of factors that influence BDNF concentrations also should be taken into account, making the assessment of BDNF concentrations pretty complex.

Although limited in scope with regard to clinical utility, our findings on between-group differences do not necessarily dismiss the possibility that abnormalities in BDNF expression reflect a pathophysiological process that may underlie depression (Duman *et al.,* 1997). Even more, the magnitude of the difference in serum BDNF concentrations between depressed patients and healthy control subjects that we report on (*e.g., d* = -0.47 for depressed persons versus healthy controls) stands out as strong when compared to other biological abnormalities in depression, for instance blood markers for immune dysregulation (*e.g.,* CRP [*d* = 0.15]) or HPA-axis activity (*e.g.,* adrenocorticotropin hormone [*d* = 0.28] see Penninx *et al.,* 2013). Furthermore, in the meta-analysis reported in **chapter 5** we find that a greater increase in serum BDNF concentrations in the course of antidepressant treatment is associated with a larger decrease in depression symptom severity. This finding may fuel work into the theoretically and clinically relevant topic on the temporal dynamics between BDNF expression and treatment efficacy. It would be interesting if future studies could address early changes in the course of (pharmacological) treatment, a notion for which some evidence exists (Lang *et al.,* 2006; Machado-Vieira *et al.,* 2009) and on which Maryna Polyakova (Max Planck Institute Leipzig, Germany) and I are writing a systematic review (work in progress).

Table 2. Alternations in serum BDNF concentrations in other psychiatric/neurological illnesses than depression (in alphabetical order). Magnitude of the difference is expressed as standardized Cohen's *d* versus a healthy control group

Disorder, author, year	Design	Finding
ADHD Corominas-Roso *et al*. (2013)	single study	BDNF concentrations are low in patients with ADHD (*n* = 54)[*d* = -0.80]
Alcohol dependency Huang *et al.* (2011)	single study	BDNF concentrations are low in alcohol dependent patients (*n* = 65)[*d* = -1.24]
Alzheimer's disease Yatsutake *et al*. (2006)	single study	BDNF concentrations are low in patients with Alzheimer's disease (*n* = 60) [*d* = -0.77]
Autism Hashimoto *et al*. (2006)	single study	BDNF concentrations are low in patients with autism (*n* = 18)[*d* = -1.58]
Bipolar Disorder Fernandes *et al.* (2011)	meta-analysis	BDNF concentrations are low in patients with bipolar disorder (*n* = 548) in the manic- and depressed state [[*d* = -0.8, 95% *CI* = -1.1 — -0.5] and *d* = -0.94, 95% *CI* = -1.72 — -0.53 respectively]
Eating disorders Monteleone *et al*. (2011)	single study	BDNF concentrations are low in patients with anorexia nervosa (*n* = 27) and bulimia (*n* = 24)[*d* = -1.54 and -1.26 respectively]
Huntington's disease Ciammola *et al*. (2007)	single study	BDNF concentrations are low in patients with Huntington's disease (*n* = 42)[*d* = -1.71]
Schizophrenia Green *et al*. (2010)	meta-analysis	BDNF concentrations are low in schizophrenic patients (*n* = 1,114)[*d* = -0.53, 95% *CI* = -0.81 — -0.18].

Abbreviations: ADHD, Attention-Hyperactivity Disorder; BDNF, Brain-Derived Neurotrophic Factor

Noteworthy in this context, and of high importance for future progress, is a recent argument by Steve Hyman (Broad Institute, Cambridge) who said in *Nature news* (May 10, 2013) that *It's a fool's errand to try to find a biomarker for a diagnosis with little basis in nature … such efforts waste human capital and governmental and industry funds.* This makes a lot of sense: the broad nosological categories that in general are used in psychiatric research pose serious limitations in the possibilities to detect (biological) abnormalities (Casey et al., 2013) because they are not valid. Searching for associations beyond the boundaries of diagnostic categories therefore may be worth considering as an important innovation. It could for instance be considered whether single, or less broad, and more carefully defined domains that may constitute the illness depression (*e.g.,* motivation or reward) in particular are correlated with BDNF. With my colleague Boudewijn Bus (Radboud University Nijmegen, The Netherlands) and others I have tried to make some advance in this (Bus *et al.,* 2013). This enterprise however yielded little additional insight. So, and given that BDNF alternations are observed in many psychopathological conditions (see **Table 2 ↑**), studies could relate alternations in BDNF to trans-diagnostic phenomena (*e.g.,* rumination; Beevers *et al.,* 2009; oxidative stress; Kapczinski *et al.,* 2008, sleep; Giese *et al.,* 2013, weight gain and loss; Monteleone *et al.,* 2007), or the research domain criteria (Casey *et al.,* 2013).

Two know more than one

Given that the above feed my concerns on the relevance of serum BDNF concentrations with regard to depression (either as a biomarker or a factor contributing to its pathophysiology) and debate on these issues in the literature I decided to ask the opinion of colleagues on this issue. Heretoo, I ran a poll in which I asked 100 researchers (who were corresponding author for published papers that had *BDNF* and *depression* (*n* = 50) or *cognitive/interpersonal* and *depression* (*n* = 50) [the latter group was included to reduce bias] in their title) about this. The results of this poll are described in detail in **Appendix V**. The majority of researchers that responded (*n* = 60) either agreed (43 percent) with the proposition that serum BDNF concentrations are relevant with regard to depression *or* expressed the belief that the future will

inform us on this issue (42 percent). Only 15 percent explicitly disagrees with the notion that serum BDNF concentrations relevant with regard to depression. In this sense, the poll was helpful in that most authors see either relevance in the use of serum BDNF concentrations as parameters for depression or suggests that more research will bring definite answers.

Recapitulating our work on serum BDNF concentrations in relation to the neurotrophin hypothesis

Table 3 ↓ gives an overview of the findings on serum BDNF concentrations and how these relate to the predictions from the neurotrophin hypothesis. Conclusions will follow in a later part of this discussion.

Table 3. Summary of the research findings in this thesis on serum BDNF concentrations and how they fit with the neurotrophin hypothesis (confrimative *versus* non-confirmative)

Confirmative
1. Serum BDNF concentrations are low in depressed patients relative to healthy controls (**chapter 4** [*N* = 923] and **chapter 5** [*N* = 5,203]).
2. Serum BDNF concentrations are normalized in the course of depression remission (**chapter 4** [*N* = 1,080] and **chapter 5** [*N* = 4,204]).
3. Serum BDNF concentrations are normalized in the course of antidepressant treatment (**chapter 4** [*N* = 1,080]; **chapter 5** [*N* = 4,204]).
4. A larger increase in serum BDNF concentrations is associated with a larger decrease in depressive symptoms over the course of antidepressant treatment (**chapter 5** [*N* = 1,422])
5. Serum BDNF concentrations are low in female patients with an anxiety disorder relative to female controls (**chapter 6** [*N* = 499]).
6. Exposure to recent stressful events is associated with lower serum BDNF concentrations (**chapter 7** [*N* = 1,435]).
Non-confirmative
1. Serum BDNF concentrations are low in the early remission phase of depression as compared to the depression state (**chapter 4** [*N* = 541])
2. Serum BDNF concentrations are normalized in the course of treatment with an antidepressant but this is not associated with remission (**chapter 4** [*N* = 421]).
3. Equally effective pharmacological antidepressants are differently associated with serum BDNF concentrations (**chapter 4** [*N* = 421]) [A]
4. Clinical characteristics (notably depression severity) are not associated with serum BDNF concentrations (**chapter 4** [*N* = 541] and **chapter 5** [*N* = 9,484]).
5. The differences in serum BDNF concentrations as a function of diagnostic and treatment status are overestimated and are of a small effect-size at best (**chapter 5** [*N* = 9,484]).
6. Serum BDNF concentrations are normal in male patients with an anxiety disorder relative to male controls (**chapter 6** [*N* = 276]).
7. Childhood trauma exposure is not associated with serum BDNF concentrations (**chapter 7** [*N* = 1,435]).

[A] Note that the possibility exist that different types of antidepressants may be clinically efficacious through different mechanisms (Mann 2005) and that therefore this finding is not neccesarly non-corfirmative with te neurotrophin hypothesis.

Besides our studies on serum BDNF this thesis explored associations between variation on the gene that codes for BDNF (val^{66}met) and depression-related phenotypes. These explorations are discussed in the following section.

BDNF val^{66}met and the neurotrophin hypothesis

The genetic studies in this thesis focused on one particular polymorphism on the BDNF gene: val^{66}met. The reason for this focus is fully described in the introduction of this thesis. In sum, the interest in val^{66}met was fuelled by two studies that showed that this polymorphism has functional properties. Egan and colleagues (2003) showed a reduced activity dependent expression of BDNF in cultured hippocampal neurons (*in vitro*) that carried a met allele. These authors extended this finding by showing worse cognitive functioning and altered hippocampal memory activity in human met carriers as compared to val/val homozygotes. In a paper published in Science (2006), Chen and colleagues confirmed these findings (*in vivo*).

Based on these groundbreaking findings, variation at the BDNF val^{66}met locus has become one of the most influential models to study BDNF functioning and it is generally believed that the field benefitted by the identification of the presumed functionality of this polymorphism (Lu *et al.,* 2013). For human studies it seems particular interesting to assess variation at the val^{66}met locus in relation to several phenotypes because it is believed that variation at this locus mirrors individual (chemical) differences in BDNF functioningin the brain. In line with this presupposition are some human studies that apparently reproduce the animal findings (*e.g.,* phenotypic hallmarks of depression such as lower hippocampal volumes in met allele carriers as compared to val/val homozygotes [Pezawas *et al.,* 2004]).

In a series of three studies, we addressed the relevance of this polymorphism with regard to: BDNF serum concentrations, DSM-IV depression and anxiety diagnoses, depression- and anxiety symptom severity, cognitive functioning, and hippocampal functioning and morphology. As an important add-on, we incorporated trauma and stress exposure in our studies to model inter-individual differences in outcomes due to these factors and their potential interaction with BDNF val^{66}met. This is imperative for the reasons that: (**I**) strong inter-individual differences exist in the degree of how detrimental the effects of trauma/stress exposure on mental health are and this may be driven by individual genetic make-up (see for example Caspi and Moffitt, 2006) and (**II**) trauma/stress exposure is a central theme in the neurotrophin hypothesis (Duman and Monteggia, 2006).

The expectations were that established correlates of depression would be related to the genotypic variant that is associated with lower neurotrophic support (i.e., the met variant) particularly in the face of trauma- or stress exposure. Some of our explorations yielded results that were in line with this expectation. Many, however, also were not. Below these findings are discussed.

BDNF val^{66}met – trauma/stress exposure and serum BDNF concentrations

In **Chapter 7** we addressed whether variation at the val^{66}met locus influences serum BDNF concentrations. The main effects of exposure to childhood abuse (*i.e.,* sexual-, physical-, or emotional abuse exposure before the age of 16 years), recent negative life events (*i.e.,* stressful events such as a divorce in the year before measurements) and their potential cross-term interactions with val^{66}met were also assesed. Our focus on the cross-term interactions among BDNF val^{66}met and stress exposure followed specifically from studies that reported that met allele carriers are more vulnerable to the effects of stress exposure as compared to individuals who are homozygous for the val allele when considering depressive symptoms (Wichers *et al*., 2008), hippocampal volume (Gatt *et al*., 2009), and cognitive functioning (Gatt *et al*., 2009).

The well-powered and controlled study reported in **Chapter 7** rendered some interesting findings. First, in the absence of main effects of trauma exposure and val^{66}met it was found that the impact of childhood abuse on serum BDNF concentrations was dependent on variation at the val^{66}met. Specifically, in met carriers, trauma exposure was associated with reduced serum BDNF concentrations, whereas in the val/val group BDNF concentrations were even higher when trauma exposure was reported (*i.e.,* a cross-over effect). The BDNF reductions that were associated with childhood abuse in met carriers were linear in nature, so that BDNF concentrations were lowest in met carriers reporting exposure to multiple types of trauma. These findings follow the neurotrophin hypothesis to some extent. Yet it should be noted that they were not in total agreement, as no associations were found between being met carrier and higher depression severity or the presence of a DSM-IV depression diagnosis, also not when exposed to childhood abuse. Maybe the conjunct of the here observed effect on serum BDNF concentrations and on psychopathology is not mandatory in order for the neurotrophin hypothesis to be valid, but it would have strenghtened the model.

A second exciting finding described in **chapter 7** was that exposure to stressful events that occurred in the past year was associated with reduced serum BDNF concentrations. This effect was, in contrast to that of childhood abuse, independent of variation at the val^{66}met locus. This result directly replicates earlier findings (*e.g.,* Trajkovksa *et al*., 2008). The finding of lower serum BDNF concentrations following stress exposure also corroborate with a body of knowledge derived from animal studies (see for instance *et al.,* 2012) and obviously also with the neurotrophin hypothesis. Interestingly, the decreased serum BDNF concentrations following recent stress exposure were, as may be expected, associated with relatively high levels of depression symptom severity. It is tempting to link these two findings, but note that these results, remarkable as they may seem, are only correlation in nature. It further should be mentioned that although statistically significant, the effect was small as it only explained ~ 1 percent of the variance in serum BDNF concentrations.

Together the findings described in **chapter 7** suggest (notwithstanding considerations as the use of cross-sectional data) that a chain of events, commencing with gene-environment interactions, may lead to low serum BDNF concentrations. It would be interesting if longitudinal studies could further unravel the developmental trajectories towards psychopathology that follow trauma and/or stress exposure and whether these may run through individual genetic make-up and neurotrophic functioning.

BDNF val^{66}met – trauma/stress exposure and the hippocampal formation

In **chapter 8** we used functional and structural MRI techniques in order to test associations between the val^{66}met variant and the structure and function of the hippocampal formation, a critical brain structure in the pathophysiology of depression (MacQueen and Frodl, 2011). For similar reasons as provided previously, we took trauma- and stress exposure into account. The study yielded the following results. First, and in line with earlier studies, we find slightly smaller hippocampal volumes in carriers of a met allele relative to val/val homozygotes. This effect has been explained as being the result of abnormal intracellular trafficking and impaired activity secretion of BDNF in carriers of a met allele (Chen *et al*., 2006). Since atrophy of the hippocampus has also been associated with (early life) stress exposure and/or a having (had) a depressive episode (MacQueen and Frodl, 2011), it is crucial to exclude the possible confounding effects of these variables. In previous studies, stress exposure and depression diagnosis have largely not been taken into account (with the exception of Frodl *et al.,* 2007 and Gatt *et al.,* 2009). We did explicitly model these interactions. It turned out, however, that the association between the met allele and lower hippocampal volume was independent of trauma/stress exposure and current/lifetime depression. This null finding is at odds with the findings reported by Gatt *et al.* (2009), who found that the combination of carrying a met allele and being exposed to early life stress was associated with particular small hippocampal volumes (and a large number of other *hallmarks* of depression such as poor cognitive functioning). The observed discrepancy between the results of Gatt *et al*. (2009) and ours may be due to a broader definition of early life stress by Gatt and colleagues (2009) who included for example also *illness* as stressful event whereas we specifically focused on childhood abuse including physical, sexual, and emotional abuseIt remains unclear how this between-study difference could have led to a different pattern of results, assuming that neither one is due to chance. With regard to the latter it should be noted that a recent large-scale study (568 healthy participants; Gerritsen *et al.,* 2012) also could not replicate the findings by Gatt *et al.* (2008). The issue of non-replication will be discussed in a next part.

In addition to on average slightly reduced hippocampal volumes, we show in **chapter 8** that val^{66}met interacts with (emotional) word valence on hippocampal encoding activity. This effect is such that hippocampal related encoding activity is increased in carriers of a met allele when presented with negative words and not when presented with neutral or positive words as compared to val/val homozygotes. This

effect was not observed in other brain areas and seems to be consistent with some studies in which emotional stimuli were used (*e.g.,* Dennis *et al*., 2010 or Lau *et al*., 2010). Although intriguing, it is imperative to mention that, as in **chapter 7** (note that the sample in **chapter 8** is a sub-sample of the much larger sample that was used in **chapter 7**), despite effects on neurobiological measures (in this case brain morphology and neuronal activity) also in this study there were no corresponding effects of the same constellation of predictor variables on psychopathology outcomes (*e.g.,* depression diagnosis, illness severity). A critical point here is that this particular study with a total *N* of only 157 may have lacked the necessary statistical power to detect between-group differences that may be small at best.In addition, with regard to the absence of associations between hippocampal volume, hippocampal function, and memory performance, a recent review showed, in line with our findings, that the model: *'a bigger brain structure → greater brain response → better performance'* may not reflect reality (Eyler *et al*., 2010).

Notwithstanding the above, the **chapters 7**, **8**, and **9** also yielded some findings that were not in line with the expectations as they can be derived from the neurotrophin hypothesis. These inconsistent findings are the topic of the section that follows.

BDNF val^{66}met – inconsistent findings

I would like to start with the finding from **chapter 9** because the study in this chapter turned out to be a defining one.

Chapter 9 reports a systematic review and meta-analysis on the association between val^{66}met and total hippocampal volume. This study was undertaken because inconsistenties have been reported with regard to this association (see for instance the difference in outcomes between Szeszko *et al.,* 2005 and Dutt *et al.,* 2009). The potential influence of demographical, clinical, and methodological characteristics of studies was also assessed. Meta-analysis confirmed that carriers of a met allele had lower hippocampal volumes relative to val/val homozygotes, yet with a very small effect-size (d = 0.13, P = .02; k = 25, total N = 3,620). However, between-study heterogeneity in effect size estimates was substantial and this could not be explained by demographical, clinical, and methodological differences across studies. Funnel plot inspection and trim-and-fill estimations suggested evidence for publication bias and effect sizes decreased substantially over the years (Pearson's correlation coefficient on the relation between year of publication and effect-size was -0.54). When publication bias was taken into account the association between val^{66}met and total hippocampal volume was no longer statistical significant. A further striking finding was that all included studies were largely underpowered. Altogether, this shows that variation at the val^{66}met locus is not likely to account for individual differences in hippocampal volume but rather that the association is subject to a *winners curse*, with large effect sizes found in a few early studies and increasingly smaller effect sizes in later (better-powered) studies.

This finding does not stand on its self. When taking a close look at the *best evidence* in the current literature a trend becomes clear. Mandelman and Grigorenko for instance (2012) pooled the data on the association between val^{66}met and general cognitive ability, memory and executive functioning (k = 23, total N = 7,095) and found, despite promising initial studies (*i.e.,* Egan *et al.,* 2003), no association between val^{66}met and cognitive functioning. Another recent meta-analysis by Kambeitz and colleagues (2013) showed, when publication bias is taken into account, that the val^{66}met polymorphism has no effect on the neuronal systems underlying the encoding of information into episodic memory (hippocampal and para-hippocampal encoding activity; 16 comparisons N = 2,985). This finding also is in contrast to what was previously claimed (e.g., Montag *et al.,* 2009). A final example is a study by Gyekis and colleageaus (2013) showing, using the largest number of subjects to date, no evidence for an association of val^{66}met with the

diagnosis depression (*k* = 26, total *N* = 17,426). This is notable since previous meta-analyses (*e.g.,* Verhagen *et al.,* 2009) have suggested that the met allele was associated with a depression diagnosis.

The pattern is clear: the evidence for associations between BDNF val^{66}met and depression relevant phenotypes is waning. Based on this I conclude that the val^{66}met polymorphism has little, if any, prediction accuracy regarding depression related phenotypes. Another consideration in formulating this conclusion is that BNDF val^{66}met was not associated (again in spite of earlier evidence) with psychopathology outcome (**chapter 7** and **8**), serum BDNF concentrations (**chapter 7**), cognitive performance (**chapter 8**), nor with hippocampal volume (**chapter 9**). So, albeit knowing the (presumed) functionality of a polymorphism (through preclinical work) the studies in this thesis show (and a large literature from other groups as well) that this is not neccesarily associated with sampled outcomes in humans. Imperative for the interpretation of the above (and also for the earlier confirmative findings) is the recent finding that many genetic variant, deletions, and copy number variants are found in neuronal cells that do not correspond with those found in non-neuronal cells (McConnel *et al.,* 2013). Therefore, our findings regarding BDNF val^{66}met, that were based on the genotype of non-neuronal cells, may be limited in that the exact correspondence between these cells and neuronal cells is not known. This issue should be acknowledged in future (single-cell) studies.

Recapitulating our work on BDNF val^{66}met in relation to the neurotrophin hypothesis

Table 4 ↓ gives an overview of the findings in this thesis that regard BDNF val^{66}met and how these relate to the predictions from the neurotrophin hypothesis. Conclusions will follow in a next part.

Table 4. Summary of the research findings in this thesis on BDNF val^{66}met and how they fit with the neurotrophin hypothesis (confrimative versus non-confirmative)

Confirmative
1. Carriers of a met allele seem to be more *vulnerable* with regard to childhood trauma exposure when serum BDNF concentrations are taken as an outcome (**chapter 7** [*N* = 1,435]).
2. Carriers of a met allele locus have somewhat lower hippocampal volumes as compared to val/val homozygotes (**chapter 8** [*N* = 157])
3. Carriers of a met allele locus show higher hippocampal activity in response to words of negative emotional valence as compared to val/val homozygotes (**chapter 8** [*N* = 157]).
Non-confirmative
1. Variation at the BDNF val^{66}met locus is not associated with serum BDNF concentrations, depression diagnosis, and depression symptom severity (**chapter 7** [*N* = 1,435]).
2. Variation at the BDNF val^{66}met locus is not associated with cognitive performance and the brain functioning (**chapter 8** [*N* = 157]).
3. Lower hippocampal volumes are not associated with carrying a met allele at the BDNF val^{66}met locus [chapter 9 [*N* = 3,620]).

The neurotrophin hypothesis and our work – recapitulating

Before I will start to contemplate on the strengths and limitations of the studies reported herein and state my conclusions, please see **Table 5 ↓** (next page) for a summary of all the findings in this thesis and how they relate to the predictions from the neurotrophin hypothesis is provided.

Table 5. Summary of research findings and how they fit with the neurotrophin hypothesis (confrimative versus non-confirmative)

Confirmative
1. Serum BDNF concentrations are low in depressed patients relative to healthy controls (**chapter 4** [*N* = 923] and **chapter 5** [*N* = 5,203]).
2. Serum BDNF concentrations are normalized in the course of depression remission (**chapter 4** [*N* = 1,080] and **chapter 5** [*N* = 4,204]).
3. Serum BDNF concentrations are normalized in the course of antidepressant treatment (**chapter 4** [*N* = 1,080]; **chapter 5** [*N* = 4,204]).
4. A larger increase in serum BDNF concentrations is associated with a larger decrease in depressive symptoms over the course of antidepressant treatment (**chapter 5** [*N* = 1,422])
5. Serum BDNF concentrations are low in female patients with an anxiety disorder relative to female controls (**chapter 6** [*N* = 499]).
6. Exposure to recent stressful events is associated with lower serum BDNF concentrations (**chapter 7** [*N* = 1,435]).
7. Carriers of a met allele seem to be more *vulnerable* with regard to childhood trauma exposure when serum BDNF concentrations are taken as an outcome (**chapter 7** [*N* = 1,435]).
8. Carriers of a met allele locus have somewhat lower hippocampal volumes as compared to val/val homozygotes (**chapter 8** [*N* = 157])
9. Carriers of a met allele locus show higher hippocampal activity in response to words of negative emotional valence as compared to val/val homozygotes (**chapter 8** [*N* = 157]).
Non-confirmative
1. Serum BDNF concentrations are low in the early remission phase of depression as compared to the active phase of depression (**chapter 4** [*N* = 541])
2. Serum BDNF concentrations are normalized in the course of treatment with an antidepressant but this is not associated with remission (**chapter 4** [*N* = 421]).
3. Several classes of equally effective pharmacological antidepressants are differently associated with serum BDNF concentrations (**chapter 4** [*N* = 421]). [A]
4. Clinical characteristics (most notably depression severity) are not associated with serum BDNF concentrations (**chapter 4** [*N* = 541]).
5. The differences in serum BDNF concentrations as a function of diagnostic and treatment status are overestimated and are of a small effect-size at best (**chapter 5** [*N* = 9,484]).
6. Serum BDNF concentrations are normal in male patients with an anxiety disorder relative to male controls (**chapter 6** [*N* = 276]).
7. Childhood trauma exposure is not associated with serum BDNF concentrations (**chapter 7** [*N* = 1,435]).
8. Variation at the BDNF val^{66}met locus is not associated with serum BDNF concentrations, depression diagnosis, and depression symptom severity (**chapter 7** [*N* = 1,435]).
9. Variation at the BDNF val^{66}met locus is not associated with cognitive performance and the brain functioning (**chapter 8** [*N* = 157]).
10. Lower hippocampal volumes are not associated with carrying a met allele at the BDNF val^{66}met locus [chapter 9 [*N* = 3,620]).

[A] Note that the possibility exist that different types of antidepressants may be clinically efficacious through different mechanisms (Mann 2005) and that therefore this finding is not neccesarly non-corfirmative with te neurotrophin hypothesis.

Conclusion

What is the final word on this? I do not think that we are on the verge of understanding depression through peripheral BDNF measurements or genetic variants that are supposed to be associated with neurotrophic functioning. The lack of universality of findings on BDNF alternations in depression that is brought forward in this thesis (and also by other research groups in recent years) suggest that attributing behavioral differences to peripheral BDNF parameters and genetic variants is overreaching. There is simply too much clinical data that do not corroborate, or are even tangential to, the predictions of the neurotrophin hypothesis. Of course, and taking for instance the heterogeneity of depression into mind, inconsistencies do not necessarily reject the neurotrophin hypothesis for all depressed patients. Besides, some predicted associations from the neurotrophic model appear to be established (*e.g.,* abnormally low serum BDNF concentrations in the depressed state). Nonetheless, in these instances the meaning of them often is not that clear (*e.g.,* reverse causation in which low neurotrophic support does not endanger a person to become depressed but rather are a consequense of being in the depressed state). In fact, reverse causation largely is my theory.

Furtermore, what this thesis illustrates clear is the value of well-powered studies, as it shows that some of the core observations on which the neurotrophin hypothesis rests are less evident, and sometimes even absent, when well-powered studies are used. This was particularly evident in our work on val^{66}met.

So, in my view the conventional wisdom that existed at the time of the start of my PhD tract that peripheral BDNF parameters and genetic variants are relevant in the pathophysiology of depression is too far fetched. In fact, from the above I conclude, whilst taking limitations into account and acknowledging that the results herein are largely contingent upon peripheral measurement that the neurotrophin hypothesis should no longer be credited in its original form.

Methodological (and other) considerations and future work

Strengths

The studies that form the heart of this thesis have salient strengths. First of all, in basically each individual study, results are derived from a large single sample or from data that come from multiple studies and together add up to a large sample. This safeguards against false positive- and negative findings and provides effect-size estimates that are accurate with regard to their magnitude (Ioannidis 2005). The proof of this principal became evident in the pooled effect-sizes that were derived from the meta-analyses that we performed, as these converged closer to those that were reported in studies that used a relatively large sample size as compared to those studies that used a relatively small sample size. A second notable strength of our work is that most analyses were adjusted for a range of possible confounding factors and that we were able to perform subgroup and moderation analyses. This allowed us to infer on (largely) independent associations, which increases, although not guarantees, the likelihood of valid findings. Validity was also achieved through the use of standardized diagnostic tools to assess current and lifetime psychopathology and the use of a control group (although the latter not necessarily eliminates all possible confounders; Prasad and Jena, 2013). Furthermore, for the interpretation of our findings we did not solely rely on *P*-values (Johnson 2013) as, where appropriate, we reported effect-size estimates and their respective confidence levels as well.

Notwithstanding strengths in design, method and reporting, I am well aware of the limitations that carry our work. The main limitations, besides those already mentioned, are discussed below.

Limitations

Table 6 ↓ lists the main limitations of the work in this thesis. These limitations are discussed in the section that follows.

Table 6. Limitations of our studies by chapter

Limitation	Because	Chapters
Cross-sectional data	Not sufficiently persuasive to prove causality	2-9 (mostly)
Non-random allocation	Not sufficiently persuasive to prove causality	2-9 (mostly)
Generalization of findings	Our study findings do not (directly) generalize to all populations	2-9 (mostly)
A multitude of tests and power	Large data-sets do not protect against multiple testing	2-9 (mostly)
	Some effects may be so small that they are hard to detect even when using a large sample	
	Some (underpowered) *post-hoc* tests may have yielded false positive- or negative findings	
Reliability, validity, and error	The measures that were used are not 100 percent reliable and valid	2-9 (mostly)

Cross-sectional data

A limitation of most of the work in this thesis is that it relies on data that were collected in a single wave. This is a limitation because data that is gathered in such a manner does not allow for conclusions that delineate the time course of event, let alone matters of causality. Take for instance our finding that serum BDNF concentrations are abnormally low in the depressed state. This finding could be explained so that a low expression of BDNF predisposes or endangers a person to an episode of depression (*i.e.,* the temporal precedence of a cause–effect relation). Indeed, this could be so. However, we cannot infer from our data that alternate explanations are false. For instance, it could be that the low serum BDNF concentrations in the depressed state are a consequence of being depressed (as has been discussed above). Note though that a lack of clarity with regard to temporal precedence is not a limitation for all cross-sectional findings that we reported. Take for instance the correlation between the amount of ambient sunshine and serum BDNF concentrations (**chapter 3**) where the presumed *cause* can be placed before the *event*, as it is not very likely that BDNF in blood causes the sun to shine. The findings regarding stress exposure and serum BDNF concentrations can be interpreted along a similar line. Yet, also for these particular cases, repeated sampling on each individual would have been more persuasive.

Non-random allocation

Another obvious limitation is that due to the epidemiological nature of the NESDA data, of which we made extensive use, none of the participants were randomly allocated to the conditions, such as medicated versus non-medicated. Therefore, our work lacks the experimental nature that is needed for causal inferences. Take for example our finding that distinct classes of antidepressants seem to have a differential effect on serum BDNF concentrations (**chapter 4**). The patients in this study were not randomly assigned to the particular antidepressant condition. Hence, *a priori* differences may have existed between persons who used a different kind of antidepressant. For instance, persons who were treated with TCA's, which is not a *first-choice* antidepressant, may have represented a clinically distinct group, consisting of a large number of non-responders on treatment with SSRI's, which typically is a *first-choice* antidepressant (Mann 2005). Thus, what we labeled as being an antidepressant-specific effect on serum BDNF concentrations may actually have been the effect of being a non-responder to treatment with a SSRI. Although we did test for a great number of possible confounders and actively explored alternative explanations, still some between-group differences may have gone undetected. Thus, because of non-random allocation, our work does not prove causality.

Generalization of findings

Other limiting factors regard the exact extent to which it can be generalized to the population at large or to specific subgroups within the population. The clinical scope of the work in this thesis is broad, as the age range of the NESDA sample is wide (18 to 65 years of age), and NESDA includes patients from several sources (*i.e.,* primary care and out-patient clinics). Notwithstanding this, in the NESDA sample no children/adolescents or elderly persons are included and the persons with depressive and anxiety disorders are all outpatients with in general modest levels of symptom severity. Also, most of the participants that were enrolled in our studies are from a Caucasian descent. So, generalizations from our findings to the young and the old, the severely ill (*e.g.,* patients who receive intra-mural care), and to persons who are not from a Caucasian background may not be straightforward. This also holds for our findings that were derived from meta-analyses, since the studies that were included also enrolled mostly persons from a Caucasian background and if they were patients, then their symptom severity was in general not that high.

Multiple testing and power

In an earlier part of this section, the use of large samples was heralded as a strong point of our work. A large sample indeed comes with advantages. However, it does not protect against the testing of a multitude of hypotheses. Given that in this thesis a substantial number of hypotheses were tested (using a single large data set) our work may have yielded some false positive findings. Second, some effects are so small that they cannot be reliably detected even when large numbers of subjects are included. The null-findings that were derived from our studies on the presumed relationship between val^{66}met and cognitive functioning (see **chapter 8**) may be a good example of this.

While considering this, statistical power is just as important to take into account when faced with positive findings (Christley 2010). Again, although overall we performed analyses using comparatively large sample sizes, at times we performed sub-group analyses that may have lacked sufficient statistical power. Likewise, the meta-regression analyses, reported in the **chapters 5** and **9**, may have been underpowered since these were based on the rather small number of studies. Our search for moderators therefore may have yielded significant associations that have different effect-sizes or actually are non-existing.

Measurement: you can never have enough precision

We measured, analyzed and concluded on BDNF concentrations in serum derived from peripheral veins. Although there are inherent advantages to this method (*i.e.,* easily accessible, only minimal invasive, and reliable with regard to intra- and inter assay variability) some points of concern should be stated.

First of all, an assumption that we had is that peripheral BDNF measurements reliably mirror the amount of BDNF in the brain. The data that underlies this assumption rely for the larger part on positive correlations between BDNF concentrations in the central nervous system and the periphery (Klein *et al*., 2010) and active transport of BDNF through the blood-brain barrier (Pan *et al*., 1998) that have been shown in non-human animal studies. Furthermore, some rodent studies have shown that peripheral administration promotes the regeneration of spinal cord injury (Krishna *et al.,* 2013) and has an effect on depressive-like behavior (Schmidt and Duman, 2010). However, there is no clear consensus on this issue and criticism and uncertainty remain. There are good reasons for this. One, in the brain, the expression of BDNF is locally and time specific (Bennet and Lagopoulos, 2013). Animal studies have shown, for instance, that antidepressant treatment increases the expression of BDNF in some brain regions (*e.g.,* the ventral tegmental area; Taliaz *et al*., 2012) but not in others (*e.g.,* the hippocampus; Lanz *et al*., 2012; Taliaz *et al*., 2012). A second reason is that there are complexities in assigning the exact sources of BDNF in peripheral tissues (Bejot *et al*., 2011). The *brain-derived* part in the name BDNF suggests that all BDNF that is active in an organism has its origin in the brain. This however is at least a little misleading (Gass and Hellweg, 2010; R Hellweg, personal communication, 2013; B Bus, personal communication, 2009 through 2013) as several types of immune-, smooth muscle-, and endothelial cells serve as sources of BDNF as well (Karege *et al*., 2002). Thus, the BDNF concentrations that we measured are likely not to have come from the brain for a 100 percent. In fact, given that serum BDNF concentrations are much higher in serum as compared to that in cerebro spinal fluid (> a 1000 fold) they may largely reflect peripheral synthesis (Pillai *et al.,* 2010). A consequence of this is that alternations in peripheral BDNF may not reflect (disturbed) central pathways but epiphenomenon of some other physiological or behavioral- and/or peripheral process that is not necessarily related to central BDNF functioning (as has been discussed above).

So, the serum BDNF measures may not more than a summed-up net, crude parameter of central BDNF functioning. Besides, some other issues regarding the measurement of serum BDNF should be acknowledged. This will be done below.

Serum BDNF concentrations versus other peripheral BDNF parameters

BDNF concentrations in serum are just one of several peripheral measures to gauge on neurotrophic functioning in the brain. Other available non-invasive options include BDNF concentrations in whole blood, blood plasma, and blood platelets. Since there are some studies that assayed a multitude of these parameters there is knowledge on how these parameters relate. In general, studies report statistically significant, yet mostly modest associations among these measures (*e.g.,* correlations between plasma and serum BDNF concentration in the range $r = 0.21$ and $r = 0.26$ (Terracciano *et al.,* 2011; Jeon *et al*., 2012) to $r = 0.71$ (Yoshimura *et al*., 2010)]. Thus, the measures that are used in the literature to gauge on the process of neurotrophic functioning in the brain relate, but far from perfect. Our findings derived from serum therefore are limited in scope in that they cannot be directly generalized to other peripheral BDNF parameters. Besides, this raises the question which parameter serves best as a mirror for neurotrophic action in the brain. Some authors have brought forward that leukocyte BDNF mRNA content, because of its short half-life, could more closely reflect central BDNF dynamics (Gass and Hellweg, 2010) and therefore perhaps may be less subject to (peripheral) confounding factors. In addition, it has been argued that a combination of peripheral BDNF indices may have advantages above a single one. Assessing both platelet and serum BDNF concentrations could be in particular relevant. Blood platelets store BDNF and release BDNF during the clotting process and by agonist simulation (Rosenfeld *et al*., 1995; Fujimura *et al*., 2002). Therefore it could be that inter-individual differences in serum BDNF concentrations are mediated by a lower activity of blood platelets caused by medications (notably here antidepressants) or pathological conditions (notably depression; Karege *et al.,* 2002).

Pro- versus mature BDNF

Besides the limitations of measuring in the periphery, there are some other drawbacks regarding the methods that we used to quantify BDNF. One is that the ELISA kit that was used in our studies could not make the distinction between the pro- and the mature BDNF variant (Lu *et al*., 2005). Thus, what we have quantified are total BDNF concentrations in serum without any regard to whether it was the pro- or the mature form. Given that the two BDNF variants are functionally different (see the introduction part for this), it would have been interesting to study whether pro-mature BDNF ratios differed, for instance, among diagnostic groups. The antibody that is sufficiently specific to make this distinction, however, was developed only recently by Yoshida *et al*. (2012a) and therefore not applied in the studies that make up this thesis.

Between-study differences and the golden standard

A final disadvantage is that large between-study differences are reported in mean serum BDNF concentrations. This has even been shown for BDNF concentrations that are assessed by the same research group, among similar diagnostic groups, using the same ELISA kit (*e.g.,* Karege *et al*., 2002 and 2005: mean serum BDNF = 22.6 ± 3.4 versus 10.1 ± 2.3 respectively). These differences probably are the result of different laboratory procedures and for within-study comparisons and meta-analyses they are not likely to constitute a limiting factor. An unfortunate consequence however is that there is no such thing as an accepted reference value that defines an individual BDNF value to be high or low. Because of this, only within-study differences can be interpreted reasonably. Standardization of measurements would be of great value here.

Summarizing the previous section, as with basically all constructs, the ability to conclude on a construct depends largely on how well the construct can be measured. Serum BDNF concentrations can be reliably measured, yet with error and with noise. Besides, the correspondence between peripheral and central

BDNF functioning is far from clear and therefore the meaning that can be assigned to (largely all) peripheral measurements is only limited.

Future work – what is worth studying and what is worth changing?

In the part that follows I discuss some options to overcome the limitations that are sketched above. These options are listed in **Table 7 ↓**.

Table 7. Areas of future interest

I. Acquire mechanistic understanding on what exactly alters neurotrophic functioning in depression
II. Single studies versus teamwork and large scale data-sharing
III. Measure and study beyond single BDNF parameters and use within-subject data
IV. Present convergent evidence from multiple research levels (*e.g.,* man and mice data in conjunct) and leave broad diagnostic categories

Mechanistic understanding

Now, and despite large interest, there is no consensus on what exactly causes altered neurotrophic functioning in depression, let alone whether it is of functional significance for (mental) health. Learning about this should be the greatest aspirational goal for the field because based on such knowledge the question whether pathological processes or epiphenomena are at play could be answered. From our studies it appears that the axiom that to depression related alternations in BDNF expression are due to trauma- or stress exposure (Duman and Monteggia, 2006) likely does not hold. Probably the relation is more complex and moderated by other factors (see **chapter 7**). Besides, there are hints on mechanisms other than stress-exposure that may regulate altered neurotrophic functioning in the depressed state. Some of these are also discussed in this thesis (*e.g.,* menopausal stage and estrogen expression). Additionally, although not empirically pinned down, I formulated two explicit hypotheses that could thrive inter-individual differences in serum BDNF concentrations in depression and in the course of treatment for this illness (*i.e.,* oxidative stress and [to depression and treatment related] changes in body-weight). Note that these hypotheses rather suggest reverse causation (*i.e.,* depression → low BDNF instead of low BDNF → depression).

Single studies and ultimate answers -- teamwork matters

In the literature it is common practice to report single study findings. For several reasons I wish to argue that the relevance of future efforts would greatly increase when other approaches were used.

Most importantly, the ever-expanding individual study results should be placed in the quantitative body of knowledge that already exists. The need for integration is bigger than ever. Data integration is important for the reason that *single studies do not provide ultimate answers*. See for instance **chapter 5** where we through data integration show that serum BDNF concentrations are not associated with the symptom severity of depression, whilst this belief initially existed (Karege *et al.,* 2002). A similar example can be found in **chapter 9** where we show, also in contrast to what was generally believed (see *e.g.,* Pezawas *et al.,* 2004), that the val^{66}met polymorphism is not associated with hippocampal volume. A data-sharing network could be the approach to answer, with rigor, many of the outstanding questions. Actually and since I truly belief in it, I am trying to launch such a project. Wide support is necessary for this, so I sought and am seeking international collaboration, with among others, Brisa Fernandes (Hospital de Clinicas, Porto Alegre, Brasil), Maryna Polyakova (Max Planck Institute, Leipzig, Germany), Kenji Hashimoto (Chiba University,

Chiba, Japan), and Rainer Hellweg (Charite University, Berlin, Germany) to come to this end. Evidently a dating-sharing network (of existing data) comes at low costs.

Beyond a single BDNF measurement

The literature to date, obviously including our own work, largely materializes on single cross-sectional measurements. Instead I would like to promote to: (**I**) measure and study beyond single BDNF measurements, (**II**) gather longitudinal (instead of between-subject) data, and (**III**) provide convergent evidence.

The common practice nowadays is to extract a single BDNF parameter from blood whilst other hormones, neurotransmitters, and receptor systems are not taken into account. This is problematic because herewith those factors that may interact with BDNF, and in theory could explain observed associations, are neglected. My eyes therefore are on studies that in conjunct to BDNF measure for instance the enzymes that convert pro-BDNF to mature BDNF or cortisol-, tryptophan-, and serotonin blood levels. A particular good example this is the recent study by Zhou *et al.* (2013) in which pro- and mature BDNF concentrations alongside their respective receptor systems: p75 and Trk-B, were assessed in serum and lymphocytes. The data from this study showed the welcome evidence that proBDNF and p75 receptors were lower in depressed patients as compared to healthy controls whereas the opposite was observed for the mature BDNF variant and its receptor Trk-B.

Next, studies should rely less on data that are collected in a single wave but instead on within-subject data. This is a more appropriate manner since it excludes a large amount of between-subject variance and an accompanying increase in the possibility to detect meaningful associations. In this thesis an example for this can be found in **chapter 5** where we show that effect-size estimates are largest when they are derived from (pure) within-subject designs. Related, it is desirable that future studies should actively control for relevant confounders (see **chapter 2** and **3**) and are sufficiently powered (see **chapter 5** and **9** for some recommendations on this).

An ideal: convergent evidence

What for me represents a general low-point in the literature is that the preclinical and clinical work on the neurotrophin hypothesis disconnect: the first uniformly reports a causal role for BDNF in the development of depressive-like behavior whereas the latter reports many null findings (including, importantly, from meta-analysis [*e.g.,* the **chapters 5** and **9**] and many other examples, *e.g.,* Kambeitz *et al.* [2012] or Dodds *et al.* [2013]). There are clashes: preclinical researchers take the stance that clinical workers do not measure the right BDNF parameters and clinical workers the one that preclinical workers do not measure depression or manipulate BDNF functioning with too much rigor (*e.g.,* completely knock it down or overexpress it a manifold of times). This leaves the neurotrophin hypothesis as a theory beyond testability. Changing cultures is necessary here.

It seems hard to directly weigh the relevance of the preclinical versus the clinical evidence. In humans the knowledge on neurotrophic functioning is largely contingent upon peripheral parameters (except maybe the studies that focus on val66met, imaging- and *post-mortem* studies). A salient detail here is that the one study that cdme closed to neurotrophic functioning in the human brain (on the amount of neuronal proliferation) by Reif *et al.* (2006) reports negative results. Preclinical studies, instead, measure BDNF in central tissues and they provide more spatial (and temporal) precision. Such studies are indispensable (albeit maybe they come with little pathological validity [Krishnan and Nestler, 2010]). It can therefore be suggested that the methods that are used in the greater part of the human studies are just not the right ones. Taking this stance should push me to reframe the title of my thesis into something like: *will the*

neurotrophin hypothesis with its predictions on depressive disorders in humans sparkle on, long after the glitter of the firework is gone? But no, this is too long of a title (Mentink A. 2013, personal communication). Besides, I am interested in human depression *per se* and there are claims that preclinical studies are too lofty and oversold (*e.g.,* manipulations that lack ecological validity) with regard to the human template they model (Couzin-Frankel 2013). Yet, progress in understanding the neurobiology of depression is contingent upon a combination of preclinical models, human cellular models and human biological studies (Hyman 2014). Besides, critical is the research evidence that comes from multiple levels (see for instance Tripp and colleagues [2013] for a good example and also that from our group, in a recent collaborative work with Maryna Polyakova (Max Planck Institute Leipzig, Germany).