Study protocol

Internet-delivered and conventional cognitive behaviour therapy for health anxiety: A randomised controlled non-inferiority trial

Summary
Health anxiety is characterized by a persistent and significant fear of having or acquiring a severe illness. The condition brings with it significant personal suffering and considerable societal costs. The most well-established treatment for health anxiety is cognitive behaviour therapy (CBT). A cost-effective way of administering CBT is via the Internet. We have demonstrated that Internet-delivered CBT (ICBT) is more efficacious than a waiting-list control condition as well as Internet-delivered behavioural stress management in reducing health anxiety. The present study aims to determine if ICBT can be non-inferior to conventional CBT – i.e., in a face-to-face format – in the treatment of health anxiety. 308 patients with health anxiety in terms of the diagnoses somatic symptom disorder or illness anxiety disorder are randomised to ICBT (n = 154) or conventional CBT (n = 154). *[Note that the design (the primary test is now based on repeated measurements) and required sample size (now 200) was revised in August 2016, see 'Health Anxiety Study - Power and Sample Size' and the text below.]* Both treatments incorporate the same interventions, primarily exposure and response prevention, and are 12 weeks long. Primary outcome measure is the Health Anxiety Inventory (HAI). Statistical analysis is based on an a priori criterion for non-inferiority (Δ). Primary hypothesis is that ICBT is at least as efficacious as conventional CBT (Δ = 0.3 Cohen's *d*). The trial is expected to take place from the autumn of 2014 to 2019.

Overview of the field (background)
Approximately five percent of the adult population suffer from *health anxiety*, or a disproportionate and persistent fear of having or acquiring a severe illness (Sunderland, Newby, & Andrews, 2013). Common examples include a strong fear of cancer, cardiovascular disease, or progressive neurological disorders (Warwick, Clark, Cobb, & Salkovskis, 1996). Health anxiety is approximately equally common in males and females, and across all age groups. About half or those afflicted develop chronic difficulties (Barsky, Fama, Bailey, & Ahern, 1998; Noyes et al., 1994).

The consequences for those afflicted are often severe. Health anxiety is not only associated with lowered health-related quality of life (Bleichhardt & Hiller, 2007), but also functional impairment in daily life activities and social interaction (Barsky et al., 1998; Noyes et al., 1993). There is also a strong associations between a strong fear of disease and psychiatric comorbidities (Gureje, Ustun, & Simon, 1997; Looper & Kirmayer, 2001; Noyes et al., 1993). Individuals with health anxiety are for example approximately 2–4 times more likely than others to meet criteria for major depressive disorder (Gureje et al., 1997; Looper & Kirmayer, 2001; Noyes, 1999).

From a societal perspective, health anxiety brings with it significant costs (Hedman et al., 2013). Compared with others, individuals with health anxiety rate their work capacity as lower and more often stay home from work (Barsky et al., 1998; Gureje et al., 1997). On average, individuals with health anxiety also consume more health care than others (Gureje et al., 1997; Looper & Kirmayer, 2001).

The aetiology of health anxiety is related to both genetic and environmental factors. The genetic contribution to health anxiety is modest (Taylor, Thordarson, Jang, & Asmundson,

2006). Retrospective correlation studies have, based on self-rated instruments, demonstrated an association between health anxiety and early experiences of being ill (Barsky, Wool, Barnett, & Cleary, 1994; Craig, Boardman, Mills, Daly-Jones, & Drake, 1993). Individuals with a strong fear of illness are also more likely than others to report psychological trauma, such as having been sexual assaulted (Barsky et al., 1994; Stein et al., 2004).

Diagnosis

In accordance with recognised diagnostic systems (DSM-IV and ICD 10), a diagnosis of *hypochondriasis* (300.7 or F45.2) often motivates the treatment of health anxiety. To meet criteria for DSM-IV hypochondriasis, a preoccupation with having or acquiring an illness need have lasted for at least 6 months. The fear is required to have persisted despite adequate medical evaluation and reassurance, and is also required to be based on misinterpretation of one's own bodily symptoms (American Psychiatric Association, 2000).

However, the *somatoform disorders* group of diagnoses, to which hypochondriasis belongs, has long been criticized. A widespread critique is that hypochondriasis is so clearly associated with worry and anxiety that the diagnosis ought to instead be classified as an anxiety disorder (Olatunji et al., 2014). Criticism has also been directed against somatic illness constituting an exclusion criterion for somatoform disorders, i.e., that the symptoms that the patients worries about need be medically unexplained if a diagnosis of DSM-IV hypochondriasis is to be considered (Rief & Martin, 2014). Given this critique, in the recently published DSM-5, the hypochondriasis diagnosis has been excluded and the diagnoses *somatic symptom disorder (SSD)* and *illness anxiety disorder (IAD)* have been introduced instead. These diagnoses allow for, but do not require, the bodily symptoms giving rise to the patient's health anxiety to be explained by a medical condition (American Psychiatric Association, 2013). For example, a patient with a well-medicated cardiovascular disease may still meet criteria for one of these diagnoses (SSD or IAD) as long as the patient's health worries are significantly excessive in relation to the actual risk of complications.

Treatment

Health anxiety can be effectively treated with both pharmacotherapy and psychological treatment. As to pharmacotherapy, studies have primarily focused on the selective serotonin reuptake inhibitors fluoxetine and paroxetine. These drugs appear to give relatively large effects (Fallon et al., 2008; Greeven et al., 2007). Side effects such as nausea, dry mouth, sexual impairment, increased perspiration, and increased anxiety are however common (Benfield, Heel, & Lewis, 1986; Greeven et al., 2007).

At the present point in time, cognitive behaviour therapy (CBT) is the psychological treatment which has the strongest evidence base for the treatment of health anxiety. Common to different forms of CBT is that the patient is expected to work independently with exercises and behaviour changes, and the treatment primarily focuses on the patient's current situation as opposed to early life experiences. CBT protocols for health anxiety may involve a series of different components, but are often based on the interventions self-observation, psychoeducation, cognitive restructuring, and exposure with response prevention (Taylor, Asmundson, & Coons, 2005).

A recently published meta-analysis was based on 13 randomised controlled trials where CBT was compared to a control condition (typically a waiting-list condition, treatment as usual, or psychological placebo) in the treatment of health anxiety. Based on aggregated between-group effect sizes, compared to the control conditions, CBT was shown to have a larger effect

on primary outcome measures, both at treatment termination (g = 0.95) and follow-up assessments (g = 0.34). Compared to the control conditions, CBT also has a larger effect on depression, both at post-treatment assessment (g = 0.64) and follow-up assessments (g = 0.35) (Olatunji et al., 2014).

Over recent years it has become increasingly common to administer cognitive behaviour therapy via the Internet. This type of CBT is best described as a form of guided self-help treatment. The patient gains access to a self-help text via the Internet, independently works with homework assignments and exercises, and communicates with his or her therapist via an email-like system. This has several advantages as compared with conventional psychological treatment. On average, therapists require less time per patient. CBT can also more readily be disseminated in rural areas and across great geographical distances. Internet-delivered CBT (ICBT) has been found to be efficacious for numerous conditions, such as major depressive disorder, social anxiety disorder, and panic disorder. The treatment form has generally shown to be as effective in routine care as in controlled trials. The effects of ICBT have often been similar to those that would be expected of conventionally administered CBT. However, only a few randomised controlled trials have conducted direct comparisons of ICBT and conventional CBT (Hedman, 2014; Hedman, Ljotsson, & Lindefors, 2012). In a recently published report from the Swedish agency for health technology assessment and assessment of social services one of the primary conclusions was that a significant limitation in the research concerning ICBT was the lack of direct comparisons between ICBT and conventional CBT (Swedish agency for health technology assessment and assessment of social services, 2013).

In a randomised controlled trial, our research group has shown that ICBT for health anxiety is more efficacious than an Internet-based waiting-list control group (Hedman, Andersson, Andersson, et al., 2011) and that it constitutes a cost-effective treatment option with persistent long-term effects at least up to one year after treatment (Hedman et al., 2013). After treatment, 67 percent of the patients receiving ICBT were in remission, which is to be compared with 2.5 percent in the control group. In a follow-up trial, we have also shown that ICBT is more efficacious than an Internet-delivered stress management program (Hedman et al., 2014). In both of these trials the within-group effects of ICBT have been large even when compared with the effects seen in conventional CBT. However, so far no direct comparison between ICBT and conventional CBT for health anxiety has been made. Before ICBT is widely implemented it is pivotal to gain more knowledge about its effects as compared with conventional CBT. This knowledge will enable more legitimate and evidence-based decisions regarding health care processes in the future.

Research questions (aims)

The aim of the present project is to investigate if Internet-delivered CBT can be at least as efficacious as – i.e., non-inferior to – conventional CBT in the treatment of health anxiety.

The main research question is: Is Internet-delivered CBT for health anxiety at least as efficacious as conventional CBT in reducing health anxiety? The primary hypothesis is that Internet-delivered CBT is at least equally efficacious as – i.e., non-inferior to – conventional CBT in reducing health anxiety, as measured on the primary outcome measure, the Health Anxiety Inventory (HAI).

Secondary research questions include the question of whether Internet-delivered CBT also is at least as efficacious as conventional CBT in reducing secondary psychiatric symptoms – i.e.,

symptoms of depression and general anxiety – as well as functional impairment. Our hypothesis is that ICBT is non-inferior to conventional CBT with regard to these secondary outcome variables.

Project description (methods)

Design and power
This is a randomised controlled trial. In total 308 patients will be recruited, 154 of these will be randomised to ICBT, and 154 will be randomised to conventional CBT. Randomisation is done in a 1:1 ratio without stratification or matching. Based on 10% attrition, 276 patients – i.e., 138 in each group – are expected to complete the post-treatment assessment. We thus expect 80% power to study effect sizes of 0.3 Cohen's *d* at alpha level .05 (Julious, 2004). *[Note that the design (the primary test is now based on repeated measurements) and required sample size (now 200) was revised in August 2016, see 'Health Anxiety Study - Power and Sample Size' and the text below.]*

Over the course of the study, as new patients are continuously recruited, these are also consecutively randomised. Patients are placed in consecutive cohorts, and randomisation is conducted for each new cohort. In a similar fashion, patients continuously begin treatment following randomisation to one of the conditions. Randomisation is done after the decision to include patients in the trial, which means that assessors are blind to forthcoming allocations to one of the two conditions in each specific case.

Recruitment
Information about the study is spread via email and letters to psychiatric clinics and primary care clinics in Stockholm County. The clinic where assessments and treatments will take place is Gustavsberg primary care clinic. If necessary, additional primary care clinics will be approached if these are willing to take part in the trial.

Application for the trial is done though self-referral via the Internet or referral from the Stockholm County routine health care system. Applicants are given access to a unique password-protected account to the study web platform. Via this account, applicants are provided patient information concerning the design of the study, prerequisites for participation, and the management of personal data. Individuals who provide informed consent complete a series of self-rated questionnaires for the purpose of screening. These questionnaires include the HAI, IAS, WI, MADRS-S, EQ-5D, AUDIT, and DUDIT (see Measurements and psychometric instruments). Applicants that have completed the screening meet with a licensed or resident psychologist to determine if participation in the study is possible, as based on the inclusion and exclusion criteria (Table 1). Applicants considered to meet criteria (i.e., all inclusion criteria but no exclusion criterion) are included as patients in the study. Those excluded are referred, if necessary, to other adequate health care clinics.

Table 1. Inclusion criteria and exclusion criteria.

Inclusion criteria	Exclusion criteria
Meet criteria for (principal) DSM-5 somatic symptom disorder or illness anxiety disorder	Other principal axis-I disorder
Registered citizen of Stockholm County	Alcohol or substance abuse or addiction in the past 6 months
At least 18 years old	Current or previous episode of psychosis or bipolar disorder

	Severe major depressive disorder
	Suicidal ideation
	Personality disorder deemed likely to severely interfere with treatment
	Initiated or non-stable continuous pharmacological treatment during the past two months, where the medication in question is deemed likely to influence the primary or secondary outcomes of the study *[Antidepressant medication]*
	Another ongoing psychological treatment for health anxiety
	Previous (in the past year) cognitive therapy or cognitive behaviour therapy for health anxiety
	Serious somatic condition that precludes CBT

Treatment

Patients undergo either ICBT for health anxiety or conventional CBT for health anxiety. The treatments are based on the same interventions and techniques, which are typical of CBT for health anxiety and have been thoroughly evaluated in our previous studies. These are primarily: information about health anxiety and CBT, systematic self-observation, exposure with response prevention (i.e., gradual confrontation with that which evokes unwanted emotional responses), and mindfulness (i.e., intentionally experiencing one's own thoughts and physical sensations without attempting to change these).

However, the treatments differ in their mode of administration. ICBT is administered through a self-help text which the patients reads via the Internet, the patient is expected to work independently with exercises and behaviour changes, and communicates with his or her therapist via an email-like system. The self-help text is segmented into 12 chapters ('modules'), each consisting of approximately 10 A4 pages of text. Each module is ended by a series of assignments based on CBT principles. The patients is encouraged to complete one module each week. Conventional CBT is instead administered through weekly face-to-face meetings between the patient and therapist. Between sessions, the patient is expected to work independently with exercises and methodical behaviour changes. One module in ICBT is comparable to one session in conventional CBT. Both treatments are 12 weeks long.

All therapists are licenced psychologists or resident psychologists with training in conducting CBT. To ensure that therapist adhere to the treatment protocol, the conventional CBT program is based on a detailed treatment manual. The trial is led by therapists with considerable experience in CBT for health anxiety, and supervision is continuously offered for all other therapists over the course of the study. Therapists also undergo initial training in the treatment manual.

Measurements and psychological instruments

To the extent that this is possible, well-established psychometric instruments are used to facilitate the psychiatric interviews of the study (Table 2). The eligibility interviews are primarily based on the Mini International Diagnostic Interview (Sheehan et al., 1998). The interview also includes other routine clinical questions (primarily regarding demographic variables, the course of symptoms over time and so on), as well as a recently developed interview for the assessment of the DSM-5 diagnoses somatic symptom disorder and illness anxiety disorder. *[Note that this refers to the Health Preoccupation Diagnostic Interview.]*

Table 2. Instruments administered by clinicians.

Abbreviation	Name	Reference

MINI	Mini International Diagnostic Interview, Swedish version 5.0.0	(Sheehan et al., 1998)
ADIS-IV	Anxiety Disorders Interview Schedule for DSM-IV (the hypochondriasis module)	(Brown, DiNardo, Barlow, & DiNardo, 1994)

Self-rated outcome measures are primarily administered before and after treatment (Table 3). Follow-up assessments are also conducted 6, as well as 12, months after treatment. The primary outcome measure of the study is the Health Anxiety Inventory (Salkovskis, Rimes, Warwick, & Clark, 2002). This is a well-established self-rated questionnaire used to evaluate treatment effects (Olatunji et al., 2014), with excellent internal consistency ($\alpha = .95$) and good test-retest-reliability ($r = .76–.90$) (Salkovskis et al., 2002).

Patients also complete a few questionnaires, primarily the Short Health Anxiety Inventory (Alberts, Hadjistavropoulos, Jones, & Sharpe, 2013), once a week over the course of the treatment. This is primarily to enable mediation analyses. In accordance with a consensus publication of the field (Rozental et al., 2014) possible adverse events are surveyed. This is done with questionnaires after treatment completion.

Table 3. Self-rated questionnaires, abbreviations and references.

Abbreviation	**Name**	**Measures**	**Reference**
HAI	Health Anxiety Inventory	Health anxiety	(Salkovskis et al., 2002)
SHAI	Short Health Anxiety Inventory	Health anxiety	(Alberts et al., 2013)
IAS	Illness Attitude Scale	Health anxiety	(Hiller, Rief, & Fichter, 2002)
WI	Whiteley Index	Health anxiety	(Pilowsky, 1967)
MADRS-S	Montgomery-Åsberg Depression Rating Scale – Self-rated version	Depression	(Svanborg & Åsberg, 1994)
BAI	Beck Anxiety Inventory	General anxiety	(Beck, Epstein, Brown, & Steer, 1988)
ASI	Anxiety Sensitivity Index	Anxiety sensitivity	(Reiss, Peterson, Gursky, & McNally, 1986)
SDS	Sheehan Disability Scale	Functional impairment	(Leon, Olfson, Portera, Farber, & Sheehan, 1997)
TIC-P	Trimbos/iMTA questionnaire for Costs associated with Psychiatric Illness	Health care consumption and productivity	(Bouwmans et al., 2013)
EQ-5D	EuroQol 5D	Health-related quality of life	(EuroQol, 1990)
WAI	Working Alliance Inventory	Therapeutic alliance	(Munder, Wilmers, Leonhart, Linster, & Barth, 2010)
C-scale	Credibility/Expectancy scale	Credibility and expectancy	(Borkovec & Nau, 1972)
CSQ-8	Client Satisfaction Questionnaire	Satisfaction with treatment	(Attkisson & Zwick, 1982)
HA-B	Health anxiety behaviours	Health anxiety behaviors	-
Flex-Val	VAS scales for the assessment of psychological flexibility	Psychological flexibility	(Hayes, Strosahl, & Wilson, 2012)
AUDIT	Alcohol Use Disorders Identification Test	Alcohol use (screening)	(Saunders, Aasland, Babor, de la Fuente, & Grant, 1993)
DUDIT	Drug Use Disorders Identification Test	Drug use (screening)	(Berman, Bergman, Palmstierna, & Schlyter, 2005)
BFI	Big Five Inventory	Personality	(John & Srivastava, 1999)

ISI	Insomnia Severity Index	Insomnia	(Morin, Belleville, Belanger, & Ivers, 2011)
OCI-R	Obsessive–Compulsive Inventory Revised	Obsessive-compulsive disorder symptoms	(Foa et al., 2002)
Y-BOCS	Yale-Brown Obsessive-Compulsive Scale	Obsessive-compulsive disorder symptoms	(Rosenfeld, Dar, Anderson, Kobak, & Greist, 1992)
SRH	Self-rated health [one-item question]	Self-rated health	(Fayers & Sprangers, 2002)
PDI-21	Peters Delusions Inventory	Psychotic symptoms	(Peters, Joseph, & Garety, 1999)
SQ	Sickness Questionnaire	Sickness/illness behaviour	(Andreasson et al., 2013)
PVD	Perceived Vulnerability for Disease	Perceived vulnerability for infectious disease	(Duncan, Schaller, & Park, 2009)
DS-R	Disgust Scale-Rev	Propensity to experience disgust	(Olatunji et al., 2007)

Data safety and data management
Psychological treatments, especially when delivered via the Internet, require safe management of personal data. Over the treatment period, all communication in ICBT is therefore done via 128-bit SSL-encrypted connections. For all study applicants that undergo the eligibility assessment a Case Report Form is safely stored. Patient records are also kept in the conventional health care system of Stockholm County (TakeCare). Analyses and combinations of patient data are only presented at a group level, in anonymised form.

Evaluation and statistical analysis
The statistical analysis is done from a non-inferiority perspective, which means that Δ, the criterion for clinically significant difference in efficacy, is determined before the trial. ICBT will be considered to be at least as efficacious as conventional CBT given that no point of the 95-percent confidence interval for the difference between the means of the treatment groups is larger than Δ to the advantage of conventional CBT. The value of Δ should ideally be based on the compilation of clinical expertise and a meta-analysis of placebo-controlled randomised controlled trials of the reference treatment (in this case conventional CBT for health anxiety) (Scott, 2009).

However, in psychological treatment research there is no clear and agreed-upon definition of what it means to control for placebo, or what it means for a trial to be ‘placebo-controlled’. The abovementioned meta-analysis (Olatunji et al., 2014) not only includes trials of conventional CBT in individual format, but also a study of ICBT and trials of CBT in group format. The article also does not differentiate between trials with different forms of control groups (e.g., waiting-list vs. treatment as usual). Nevertheless, the primary result from the study, in particular in combination with the Olatunji and colleagues graphical presentation of study effects, is likely to give an approximate estimate of how much better conventional CBT is, compared to a psychological treatment without a specific treatment effect. The lowest point of the confidence interval for the controlled effect on the primary outcome of the trials was $g = 0.66$ (Olatunji et al., 2014). As to secondary outcome measures, based on 95-percent confidence intervals, the meta-analysis presented the lowest value of $g = 0.35$ for general anxiety and the lowest value of $g = 0.41$ for symptoms of depression (Olatunji et al., 2014). Given that these analyses were based on several comparisons against waiting-lists effects are probably larger than those that it would be reasonable to consider ‘placebo-controlled’.

In the present trial, the non-inferiority margin (Δ) is set to 0.3 Cohen’s *d*. This is a reasonable upper limit for how much more efficacious conventional CBT can be allowed to be in relation

to ICBT without ICBT being considered as clinically significantly less efficacious. Based on data from pervious trials (Hedman, Andersson, Andersson, et al., 2011; Hedman et al., 2014) we expect 0.3 Cohen's *d* to approximate 7.5 points on the primary outcome (i.e., the Health Anxiety Inventory [HAI]).

[Note that the design and required sample size was revised on August 31st 2016, when 126 patients had been randomised and no data had been analysed. As stated above, we originally powered the trial for a post-treatment mean difference test of non-inferiority, and planned for a sample size of 308. In the revised design, we decided to instead base the primary analysis on the weekly health anxiety assessments (i.e., the more commonly used version of the Health Anxiety Inventory that is here referred to as the 'SHAI'), as considerably higher precision would be achieved. Based on Monte Carlo simulation and a linear mixed modelling approach, a sample size of 200 was estimated to be sufficient to confirm non-inferiority with 80% power, given a true between-group effect size of zero, and the expected pattern of data loss. See 'Health Anxiety Study - Power and Sample Size' below for further details.]

For the primary outcome measure, the Health Anxiety Inventory (HAI), and secondary self-rated scales, group means with 95-percent confidence intervals for the pre- and post-treatment assessments are presented. Effect sizes are calculated based on Cohen's *d*. The primary hypothesis test of non-inferiority (see above) is based on per-protocol data, but values of M and *d* based on estimated values may also be presented given that these are deemed to be informative. Additional significance tests are based on mixed effects modelling and the intention-to-treat principle which implies that all data from all patients are included in the analyses. Clinically significant improvement will also be analysed based on the Jacobson and Truax algorithm (Jacobson & Truax, 1991). Nominal data are primarily analysed based on the χ^2-test. *[Note that the design and required sample size was revised in August 2016, see above, and 'Health Anxiety Study - Power and Sample Size' below.]*

Time plan and feasibility

The study is initiated at the end of 2014 and is estimated to continue until the end of 2018 or early 2019, i.e., for approximately 4 years. First, a preparation period will take approximately 1 month. This period is used to develop treatment manuals and procedures for recruitment and other key aspects of the trial. Thereafter follows a pilot period of approximately 1.5–2 months, with the first 8 patients of the study. After this, 100 patients will be recruited per year, each year in 4 cohorts (26+24+26+24), over the course of 3 years. Recruitment and treatment will take place in parallel, so that the next cohort is recruited as the pervious is assigned to treatment:

Approx. week 2–13	Approx. week 14-25	Approx. week 30-41	Approx. week 41-52
Treatment cohort 1	Treatment cohort 2	Treatment cohort 3	Treatment cohort 4
Recruitment cohort 2	Recruitment cohort 3	Recruitment cohort 4	Recruitment cohort 5

Until the end of 2015 we thus expect to treat (8 + 100 =) 108 persons. Until the end of 2016 a total of (8 + 2·100 =) 208 persons. Around the end of 2017 or early 2018 we expect 308 persons to have received treatment. The last 12-months follow up will be conducted around the end of 2018 or early 2019.

Our assessment is that the necessary requirements to complete this project according to plan are met. To a large extent, the structural requirements for the procedure have already been established. Treatment manuals will be easily based on the self-help content evaluated in previous trials. Facilities and physical resources are readily available through Gustavsberg primary care clinic. We also have access to a well-established web-based treatment platform, originally developed for the Internet psychiatry unit of Psychiatry Southwest.

Research group
Researchers involved have documented extensive experience in the treatment of health anxiety and our research group has conducted numerous randomised controlled trials of CBT (Hedman et al., 2012). The principal investigator of the trial (Erik Hedman, lic. psychologist, PhD) has conducted several previous trials of CBT for health anxiety (Hedman, Andersson, Andersson, et al., 2011; Hedman et al., 2014; Hedman et al., 2010), social anxiety disorder (Furmark et al., 2009; Hedman, Andersson, Ljotsson, et al., 2011), and irritable bowel syndrome (Ljótsson et al., 2010; Ljótsson et al., 2011). The project leader of the trial (Erland Axelsson, MSc) has assisted in two of Erik Hedman's previous trials CBT for health anxiety. Erland has conducted more than 400 diagnostic interviews focused on health anxiety. He has also conducted approximately 35 Internet-delivered treatments for health anxiety. All therapists at Gustavsberg primary care clinic have documented training in cognitive behaviour therapy. Over the course of the trial, if needed to assess the trial, the research group also involves other experts such as Brjánn Ljótsson (lic. psychologist, PhD) with experience in psychological treatment research and cognitive behaviour therapy for numerous conditions such as health anxiety (Hedman, Andersson, Andersson, et al., 2011), fibromyalgia (Ljótsson et al., 2014), and irritable bowel syndrome (Ljótsson et al., 2010; Ljótsson et al., 2011).

Implications
The present study is expected to, if completed, be the first to directly compare ICBT and conventional CBT in the treatment of health anxiety. Based on the observed effects the study would provide new evidence regarding to what extent these two treatment dissemination formats (that is, Internet-delivered treatment versus face-to-face treatment) should be prioritised and best utilised to treat health anxiety in the routine health care environment. If, for example, ICBT is shown to give effects on par with those of conventional CBT this will provide additional incentives to make ICBT for health anxiety more widely available.

The study would also be one of very few direct comparisons of ICBT and conventional CBT *per se* (regardless of patient group). Thus, the results of the trial would be of great value also for a more overarching discussion concerning the value of Internet-delivered psychological treatment in relation to conventional alternatives.

References

Alberts, N. M., Hadjistavropoulos, H. D., Jones, S. L., & Sharpe, D. (2013). The Short Health Anxiety Inventory: A systematic review and meta-analysis. *Journal of Anxiety Disorders, 27*(1), 68-78. doi: http://dx.doi.org/10.1016/j.janxdis.2012.10.009

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders : DSM-5*. Arlington, Va.: American Psychiatric Association.

Andreasson, A. N., Wicksell, R., Karshikoff, B., Lodin, K., Axelsson, J., Olgart Höglund, C., & Lekander, M. (2013). 49. Development and preliminary validation of the Sickness Questionnaire (SicknessQ). *Brain, Behavior, and Immunity, 32*, e14.

Attkisson, C. C., & Zwick, R. (1982). The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and program planning, 5*(3), 233-237.

Barsky, A. J., Fama, J. M., Bailey, E. D., & Ahern, D. K. (1998). A prospective 4- to 5-year study of DSM-III-R hypochondriasis. *Arch Gen Psychiatry, 55*(8), 737-744.

Barsky, A. J., Wool, C., Barnett, M. C., & Cleary, P. D. (1994). Histories of childhood trauma in adult hypochondriacal patients. *The American Journal of Psychiatry, 151*(3), 397-401.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of consulting and clinical psychology, 56*(6), 893.

Benfield, P., Heel, R. C., & Lewis, S. P. (1986). Fluoxetine. A review of its pharmacodynamic and pharmacokinetic properties, and therapeutic efficacy in depressive illness. *Drugs, 32*(6), 481-508.

Berman, A. H., Bergman, H., Palmstierna, T., & Schlyter, F. (2005). Evaluation of the Drug Use Disorders Identification Test (DUDIT) in criminal justice and detoxification settings and in a Swedish population sample. *Eur Addict Res, 11*(1), 22-31. doi: 10.1159/000081413

Bleichhardt, G., & Hiller, W. (2007). Hypochondriasis and health anxiety in the German population. *British Journal of Health Psychology, 12*(4), 511-523. doi: http://dx.doi.org/10.1348/135910706X146034

Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry, 3*(4), 257-260. doi: http://dx.doi.org/10.1016/0005-7916%2872%2990045-6

Bouwmans, C., De Jong, K., Timman, R., Zijlstra-Vlasveld, M., Van der Feltz-Cornelis, C., Tan Swan, S., & Hakkaart-van Roijen, L. (2013). Feasibility, reliability and validity of a questionnaire on healthcare consumption and productivity loss in patients with a psychiatric disorder (TiC-P). *BMC Health Serv Res, 13*, 217. doi: 10.1186/1472-6963-13-217

Brown, T. A., DiNardo, P., Barlow, D. H., & DiNardo, P. A. (1994). *Anxiety Disorders Interview Schedule for DSM-IV.: Client interview schedule*: Oxford University Press.

Craig, T., Boardman, A., Mills, K., Daly-Jones, O., & Drake, H. (1993). The South London Somatisation Study: I. Longitudinal course and the influence of early life experiences. *The British Journal of Psychiatry, 163*, 579-588. doi: http://dx.doi.org/10.1192/bjp.163.5.579

Duncan, L. A., Schaller, M., & Park, J. H. (2009). Perceived vulnerability to disease: Development and validation of a 15-item self-report instrument. *Personality and Individual differences, 47*(6), 541-546.

EuroQol, G. (1990). EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy, 16*(3), 199-208.

Fallon, B. A., Petkova, E., Skritskaya, N., Sanchez-Lacay, A., Schneier, F., Vermes, D., . . . Liebowitz, M. R. (2008). A double-masked, placebo-controlled study, of fluoxetine for hypochondriasis. *Journal of Clinical Psychopharmacology, 28*(6), 638-645. doi: http://dx.doi.org/10.1097/JCP.0b013e31818d21cf

Fayers, P. M., & Sprangers, M. A. (2002). Understanding self-rated health. *Lancet, 359*(9302), 187-188. doi: 10.1016/s0140-6736(02)07466-4

Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive-Compulsive Inventory: development and validation of a short version. *Psychol Assess, 14*(4), 485-496.

Furmark, T., Carlbring, P., Hedman, E., Sonnenstein, A., Clevberger, P., Bohman, B., . . . Andersson, G. (2009). Guided and unguided self-help for social anxiety disorder: randomised controlled trial. *Br J Psychiatry, 195*(5), 440-447. doi: 10.1192/bjp.bp.108.060996

Greeven, A., van Balkom, A. J., Visser, S., Merkelbach, J. W., van Rood, Y. R., van Dyck, R., . . . Spinhoven, P. (2007). Cognitive behavior therapy and paroxetine in the treatment of hypochondriasis: A randomized controlled trial. *The American Journal of Psychiatry, 164*(1), 91-99. doi: http://dx.doi.org/10.1176/appi.ajp.164.1.91

Gureje, O., Ustun, T., & Simon, G. (1997). The syndrome of hypochondriasis: A cross-national study in primary care. *Psychological Medicine, 27*(5), 1001-1010. doi: http://dx.doi.org/10.1017/S0033291797005345

Hayes, S., Strosahl, K., & Wilson, K. (2012). *Acceptance and Commitment Therapy: The Process and Practice of Mindful Change* (2 ed.). New York, NY: The Guilford Press.

Hedman, E. (2014). Therapist guided internet delivered cognitive behavioural therapy. *BMJ, 348*, g1977. doi: 10.1136/bmj.g1977

Hedman, E., Andersson, E., Lindefors, N., Andersson, G., Ruck, C., & Ljotsson, B. (2013). Cost-effectiveness and long-term effectiveness of Internet-based cognitive behaviour therapy for severe health anxiety. *Psychological Medicine, 43*(2), 363-374. doi: http://dx.doi.org/10.1017/S0033291712001079

Hedman, E., Andersson, G., Andersson, E., Ljotsson, B., Ruck, C., Asmundson, G. J., & Lindefors, N. (2011). Internet-based cognitive-behavioural therapy for severe health anxiety: randomised controlled trial. *Br J Psychiatry, 198*(3), 230-236. doi: 10.1192/bjp.bp.110.086843

Hedman, E., Andersson, G., Ljotsson, B., Andersson, E., Ruck, C., Mortberg, E., & Lindefors, N. (2011). Internet-based cognitive behavior therapy vs. cognitive behavioral group therapy for social anxiety disorder: a randomized controlled non-inferiority trial. *PLoS One, 6*(3), e18001. doi: 10.1371/journal.pone.0018001

Hedman, E., Axelsson, E., Gorling, A., Ritzman, C., Ronnheden, M., El Alaoui, S., . . . Ljotsson, B. (2014). Internet-delivered exposure-based cognitive-behavioural therapy and behavioural stress management for severe health anxiety: randomised controlled trial. *Br J Psychiatry*. doi: 10.1192/bjp.bp.113.140913

Hedman, E., Ljotsson, B., Andersson, E., Ruck, C., Andersson, G., & Lindefors, N. (2010). Effectiveness and cost offset analysis of group CBT for hypochondriasis delivered in a psychiatric setting: An open trial. *Cognitive Behaviour Therapy, 39*(4), 239-250. doi: http://dx.doi.org/10.1080/16506073.2010.496460

Hedman, E., Ljotsson, B., & Lindefors, N. (2012). Cognitive behavior therapy via the Internet: a systematic review of applications, clinical efficacy and cost-effectiveness. *Expert Rev Pharmacoecon Outcomes Res, 12*(6), 745-764. doi: 10.1586/erp.12.67

Hiller, W., Rief, W., & Fichter, M. (2002). Dimensional and categorical approaches to hypochondriasis. *Psychological Medicine, 32*(4), 707-718. doi: http://dx.doi.org/10.1017/S0033291702005524

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol, 59*(1), 12-19.

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research, 2*(1999), 102-138.

Julious, S. A. (2004). Sample sizes for clinical trials with normal data. *Stat Med, 23*(12), 1921-1986. doi: 10.1002/sim.1783

Leon, A. C., Olfson, M., Portera, L., Farber, L., & Sheehan, D. V. (1997). Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. *Int J Psychiatry Med, 27*(2), 93-105.

Ljótsson, B., Atterlof, E., Lagerlof, M., Andersson, E., Jernelov, S., Hedman, E., . . . Wicksell, R. K. (2014). Internet-delivered acceptance and values-based exposure treatment for fibromyalgia: a pilot study. *Cogn Behav Ther, 43*(2), 93-104. doi: 10.1080/16506073.2013.846401

Ljótsson, B., Falk, L., Vesterlund, A. W., Hedman, E., Lindfors, P., Ruck, C., . . . Andersson, G. (2010). Internet-delivered exposure and mindfulness based therapy for irritable bowel syndrome--a randomized controlled trial. *Behav Res Ther, 48*(6), 531-539. doi: 10.1016/j.brat.2010.03.003

Ljótsson, B., Hedman, E., Andersson, E., Hesser, H., Lindfors, P., Hursti, T., . . . Andersson, G. (2011). Internet-delivered exposure-based treatment vs. stress management for irritable bowel syndrome: a randomized trial. *The American journal of gastroenterology, 106*(8), 1481-1491.

Looper, K. J., & Kirmayer, L. J. (2001). Hypochondriacal concerns in a community population. *Psychological Medicine, 31*(4), 577-584. doi: http://dx.doi.org/10.1017/S0033291701003737

Morin, C. M., Belleville, G., Belanger, L., & Ivers, H. (2011). The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep, 34*(5), 601-608.

Munder, T., Wilmers, F., Leonhart, R., Linster, H. W., & Barth, J. (2010). Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. *Clin Psychol Psychother, 17*(3), 231-239. doi: 10.1002/cpp.658

Noyes, R. (1999). The relationship of hypochondriasis to anxiety disorders. *General Hospital Psychiatry, 21*(1), 8-17.

Noyes, R., Kathol, R. G., Fisher, M. M., Phillips, B. M., Suelzer, M. T., & Holt, C. S. (1993). The validity of DSM-III-R hypochondriasis. *Archives of General Psychiatry, 50*(12), 961-970. doi: http://dx.doi.org/10.1001/archpsyc.1993.01820240045006

Noyes, R., Kathol, R. G., Fisher, M. M., Phillips, B. M., Suelzer, M. T., & Woodman, C. L. (1994). One-year follow-up of medical outpatients with hypochondriasis. *Psychosomatics: Journal of Consultation Liaison Psychiatry, 35*(6), 533-545. doi: http://dx.doi.org/10.1016/S0033-3182%2894%2971722-9

Olatunji, B. O., Kauffman, B. Y., Meltzer, S., Davis, M. L., Smits, J. A., & Powers, M. B. (2014). Cognitive-behavioral therapy for hypochondriasis/health anxiety: A meta-analysis of treatment outcome and moderators. *Behav Res Ther, 58*, 65-74. doi: 10.1016/j.brat.2014.05.002

Olatunji, B. O., Williams, N. L., Tolin, D. F., Abramowitz, J. S., Sawchuk, C. N., Lohr, J. M., & Elwood, L. S. (2007). The Disgust Scale: item analysis, factor structure, and suggestions for refinement. *Psychol Assess, 19*(3), 281-297. doi: 10.1037/1040-3590.19.3.281

Peters, E. R., Joseph, S. A., & Garety, P. A. (1999). Measurement of delusional ideation in the normal population: introducing the PDI (Peters et al. Delusions Inventory). *Schizophr Bull, 25*(3), 553-576.

Pilowsky, I. (1967). Dimensions of hypochondriasis. *Br J Psychiatry, 113*(494), 89-93.

Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behav Res Ther, 24*(1), 1-8.

Rief, W., & Martin, A. (2014). How to use the new DSM-5 somatic symptom disorder diagnosis in research and practice: a critical evaluation and a proposal for modifications. *Annu Rev Clin Psychol, 10*, 339-367. doi: 10.1146/annurev-clinpsy-032813-153745

Rosenfeld, R., Dar, R., Anderson, D., Kobak, K. A., & Greist, J. H. (1992). A computer-administered version of the Yale-Brown Obsessive-Compulsive Scale. *Psychological Assessment, 4*(3), 329.

Rozental, A., Andersson, G., Boettcher, J., Ebert, D. D., Cuijpers, P., Knaevelsrud, C., . . . Carlbring, P. (2014). Consensus statement on defining and measuring negative effects of Internet interventions. *Internet Interventions, 1*(1), 12-19.

Salkovskis, P. M., Rimes, K., Warwick, H., & Clark, D. (2002). The Health Anxiety Inventory: Development and validation of scales for the measurement of health anxiety and hypochondriasis. *Psychological Medicine, 32*(5), 843-853. doi: http://dx.doi.org/10.1017/S0033291702005822

Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. *Addiction, 88*(6), 791-804.

Scott, I. A. (2009). Non-inferiority trials: determining whether alternative treatments are good enough. *Med J Aust, 190*(6), 326-330.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., . . . Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry, 59 Suppl 20*, 22-33;quiz 34-57.

Stein, M. B., Lang, A. J., Laffaye, C., Satz, L. E., Lenox, R. J., & Dresselhaus, T. R. (2004). Relationship of sexual assault history to somatic symptoms and health anxiety in women. *General Hospital Psychiatry, 26*(3), 178-183. doi: http://dx.doi.org/10.1016/j.genhosppsych.2003.11.003

Sunderland, M., Newby, J. M., & Andrews, G. (2013). Health anxiety in Australia: Prevalence, comorbidity, disability and service use. *The British Journal of Psychiatry, 202*(1), 56-61. doi: http://dx.doi.org/10.1192/bjp.bp.111.103960

Svanborg, P., & Åsberg, M. (1994). A new self-rating scale for depression and anxiety states based on the Comprehensive Psychopathological Rating Scale. *Acta Psychiatrica Scandinavica, 89*(1), 21-28.

Swedish agency for health technology assessment and assessment of social services. (2013). Internetförmedlad psykologisk behandling vid ångest- och förstämningssyndrom [Internet-delivered psychological treatment in anxiety and mood disorders]. *SBU Alert-rapport* (Vol. 2013-02).

Taylor, S., Asmundson, G. J., & Coons, M. J. (2005). Current directions in the treatment of hypochondriasis. *Journal of Cognitive Psychotherapy, 19*(3), 285-304.

Taylor, S., Thordarson, D. S., Jang, K. L., & Asmundson, G. J. (2006). Genetic and environmental origins of health anxiety: A twin study. *World Psychiatry, 5*(1), 47.

Warwick, H. M., Clark, D. M., Cobb, A. M., & Salkovskis, P. M. (1996). A controlled trial of cognitive-behavioural treatment of hypochondriasis. *Br J Psychiatry, 169*(2), 189-195.

Date: 2014-12-02 *[December 2nd 2014]*

To: The regional ethics review board of Stockholm

Amendment to approved application with id 2014/1530-31/2

Principal investigator

Erik Hedman, PhD, lic. psychologist, Karolinska Institutet, the Department of Clinical Neuroscience, and Stockholm County Council.

Background

Our research group has been given ethics approval (id 2014/1530-31/2) to, by means of a randomised controlled trial, compare the effects of Internet-based cognitive behaviour therapy with the effects of conventional cognitive behaviour therapy in the treatment of health anxiety. The primary outcome measure of the study is health anxiety, but other symptom domains are also investigated.

Amendment – Assessment of functional impairment with the WHODAS 2.0

An important aspect of ill mental health is functional impairment, or an inability to engage in, and meet the requirements of, daily life.

The diagnostic system Diagnostic and Statistical Manual of Mental disorders (DSM) has recently been published in a new version (DSM-5) (1). The DSM-5 recommends usage of the World health organization (WHO) instrument WHO Disability Assessment Schedule (WHODAS) 2.0 in the assessment of functional impairment (2). The Swedish National Board of Health and Welfare plans to publish an authorized Swedish version of the instrument in December 2014 (3).

Our research group therefore wishes to make an amendment to the approved ethics application. This amendment is that **the short self-rated version of the WHODAS 2.0 is added as a secondary outcome measure of the trial.**

[This copy is not signed since it is an English translation of the Swedish original.]

_____

Erik Hedman, lic. psychologist, PhD, principal investigator of the trial
kire.hedman@ki.se; telephone, 0709-66 70 74

References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA.: American Psychiatric Publishing; 2013.
2. World Health Organization. WHO Disability Assessment Schedule 2.0 (WHODAS 2.0) 2014 [Available from: http://www.who.int/classifications/icf/whodasii/en/.
3. **Persson KB**. Conversation with the Swedish National Board of Health and Welfare via email. 2014-10-20.

Health Anxiety Study - Power and Sample Size

Michael C Sachs

2016-08-31

Summary

To analyze the type of trial under consideration, we recommend a linear mixed effects model with the treatment group by visit time interaction being the main parameter of interest. A model with random intercepts and slopes would suffice to account for the correlated observations and missing data. The treatment by time interaction, multiplied by 12, can be interpreted as the average difference between treatment groups in the change in HAI score over 12 weeks. For instance, in the previous study, we might report that over the course of 12 weeks of therapy, BSM treatment led to a 2.4 point smaller decrease in HAI score as compared to KBI (95% CI: 0.1 to 4.6 points). Report the changes over 12 weeks in each treatment group with confidence intervals for ease of interpretation.

In the new study, for a sample size of 250 [*Erratum. The proper text should read "200" for a one-sided test; see page 7–i.e., the Results section–of this power analysis report. /Principal investigator*] individuals total, and a true treatment difference of 0 d, there is approximately 80% power to confirm noninferiority, using a 95% confidence interval to rule out the margin of 0.3 Cohen's d.

Background and Assumptions

The investigators are planning a randomized clinical trial comparing two modalities of cognitive behavioral therapy (CBT): internet-based self-guided and face-to-face. The treatments each last for 12 weeks. The primary outcome measure is the short version of the health anxiety inventory (HAI). The HAI will be measured at baseline before treatment (time 0) and 12 more times after that at the end of each week.

The main hypothesis is that internet-based CBT is no worse at improving health related anxiety compared to face to face CBT. In previous trials, face-to-face CBT has been shown to be efficacious at decreasing the HAI as compared to a control treatment. Several studies report an effect size of 1.4 - 1.5 (Cohen's d), while one study showed a much smaller effect size of 0.3 - 0.4.

The trial under consideration here is therefore planned as a non-inferiority study, with a noninferiority margin of 0.3. That is, we aim to demonstrate conclusively that internet-based CBT is no more than 0.3 standard deviations (20% worse, assuming an effect of

1.5) worse than face-to-face CBT in terms of reducing the HAI over the course of 12 weeks. This margin was determined based on clinical judgement in consideration of the previously reported effect sizes in similar studies.

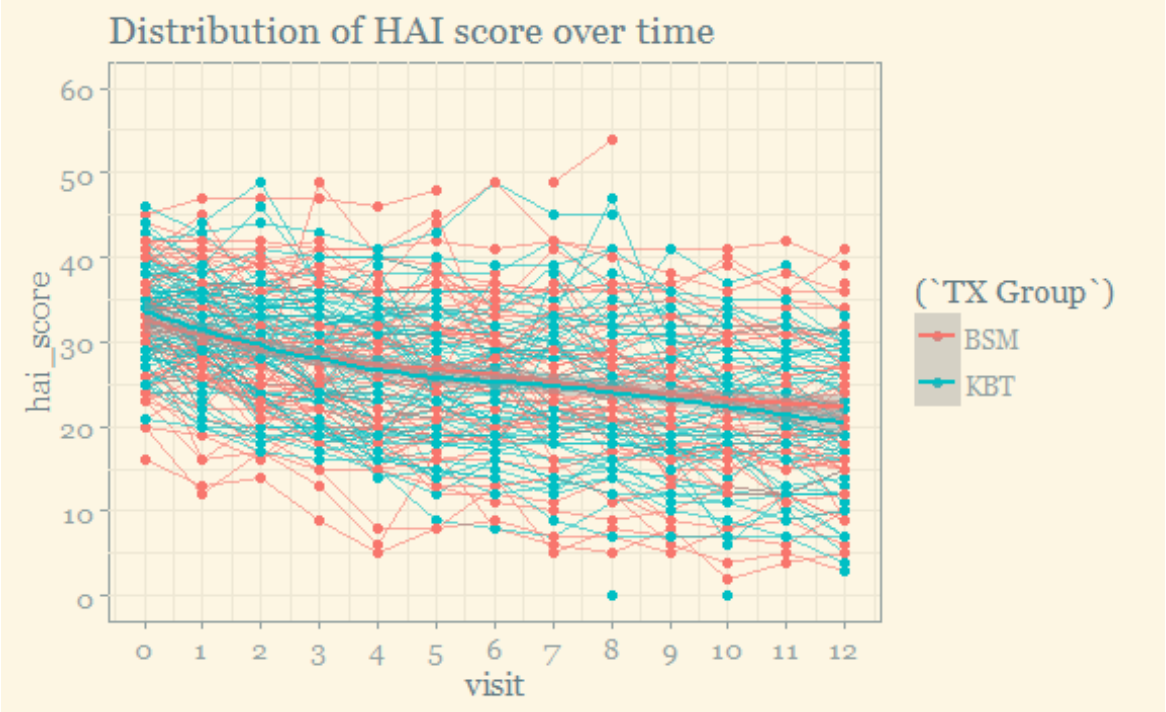
Based on preliminary data, the average change in HAI score over 12 weeks was approximately -10 points, and the standard deviation of that change was about 7.5 points. Therefore, a Cohen's d of 0.3 corresponds to a difference in slopes of 7.5 * 0.3 = 2.25. Thus 2.25 will be our non-inferiority margin on the scale of changes over 12 weeks.

Data Analysis and Interpretation

Previous studies have only compared the post treatment HAI score between groups. There is a wealth of information contained in the other 12 weeks worth of data that can be used to improved the statistical efficiency of the analysis. Here we will demonstrate how using data from a previous study.

Preliminary data used to obtain estimates for key parameters we obtained from the file called "Health anxiety example data to Michael with password.xlsx" sent on 2016-08-19 from Erik Hedman. This trial is similar to the one being planned. We use these data to guide our assumptions on the parameters used to calculated power for the new trial.

The following plot shows the distributions of the HAI scores over time, by treatment group. The smooth curves are flexible fitted models that represent the average trend in each group. The trends are roughly linear over the observation period.



A simple way to analyze this trial would be to compare the mean HAI score at the end of the study by treatment group. Since the trial is randomized, we know that there is no true difference in HAI score at baseline. However, due to sampling variability, there may be an observed difference in mean HAI score at baseline between treatment groups. Furthermore, since the distribution of HAI score at baseline has a fairly wide spread, we can gain efficiency by calculating the change in HAI score from baseline to 12 weeks for each subject, and then comparing the mean change in HAI score by treatment group.

A logical extension to that is to use information from each visit to estimate the average change in HAI score for each subject. This is called a derived variable analysis, where for each subject, the derived variable is the slope for the change in HAI score over time. We then compare the average slopes by treatment group.

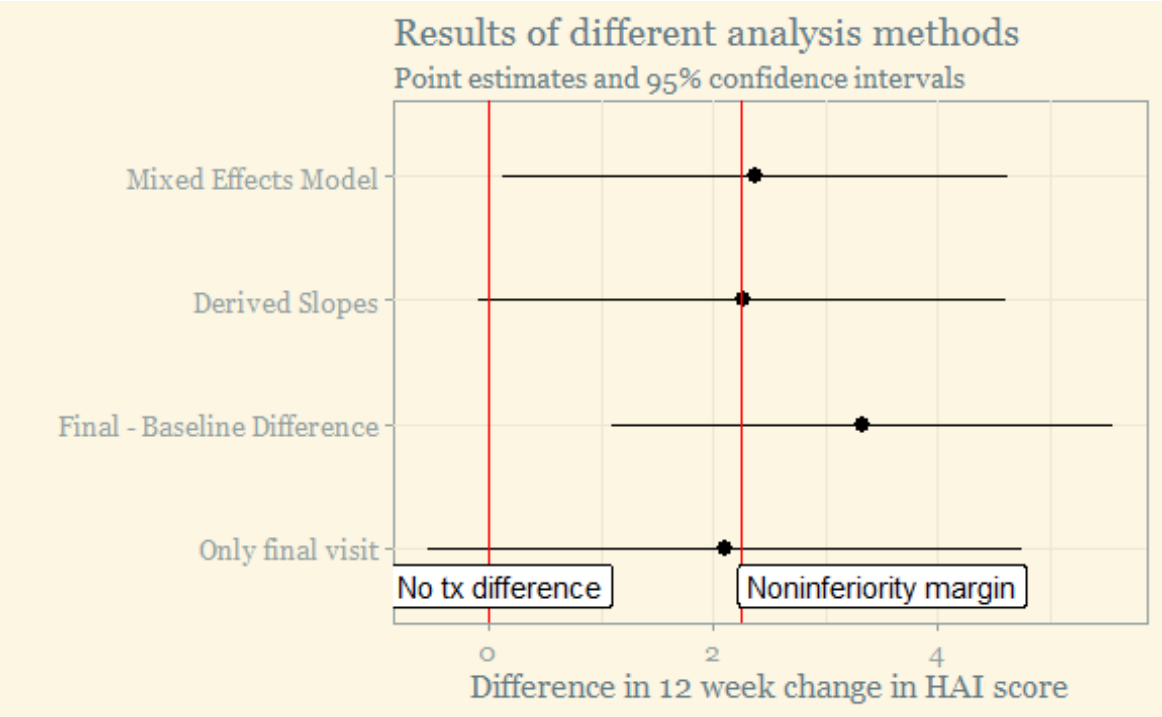
A further extension to this concept is to use linear mixed effects models to estimate the average difference between treatment groups in the change in HAI score over time. Such a model would account for the within-subject correlation over time, in addition to the intermittent missing data. Specifically, the model is

$$Y_{it} = (\alpha + a_i) + (\beta + b_i) * t + \delta * Z_i + \gamma * t * Z_i + \varepsilon_i,$$

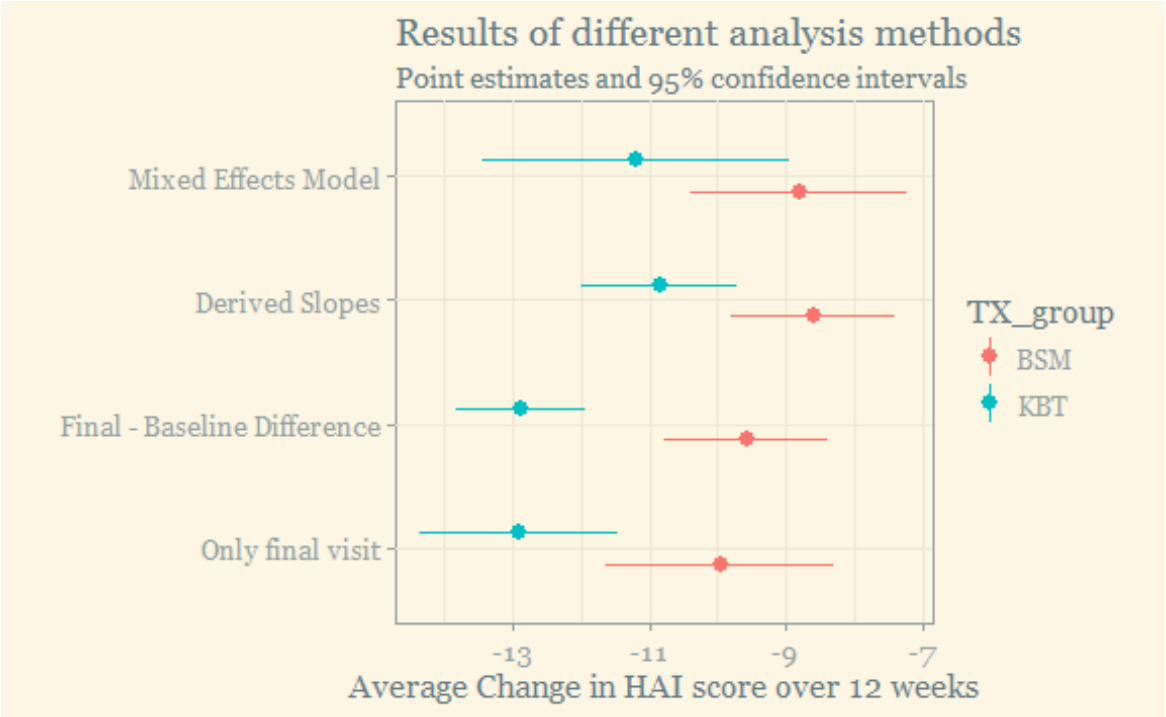
where Y_{it} is the HAI score for subject i at visit t, Z_i is the treatment group for subject i, and a_i, b_i, ε_i are normally distributed random variables with mean 0 (random effects). This model can be referred to as a linear mixed effects model with random intercepts and random slopes. The fixed effects are parameters for the time effect, the treatment effect, and their interaction. The parameter of interest is the interaction parameter γ, which represents the average difference between treatment groups in the change in HAI score per week. The model can be estimated with maximum likelihood using a variety of software. One example using the lme4 package in R is given in the appendix. We also provide a link to the SPSS documentation for mixed effects models.

The results of these different analysis methods are shown in the next figure. The parameters shown all have the same interpretation: the difference between treatment groups in the standard deviation change in HAI score after 12 weeks of therapy. In this study all confidence intervals overlap with the noninferiority margin. The final visit and derived slopes analysis confidence intervals also overlap with 0, meaning that those results are inconclusive. The other two confidence intervals exclude 0, indicating that one treatment is superior.

The width of the confidence intervals is an indicator of the precision of the parameter estimates. The confidence interval width of the "Only final visit" method is 5.3, compared to the others that range from 4.45 to 4.7. This supports our belief that the "Only final visit" approach is clearly inferior.



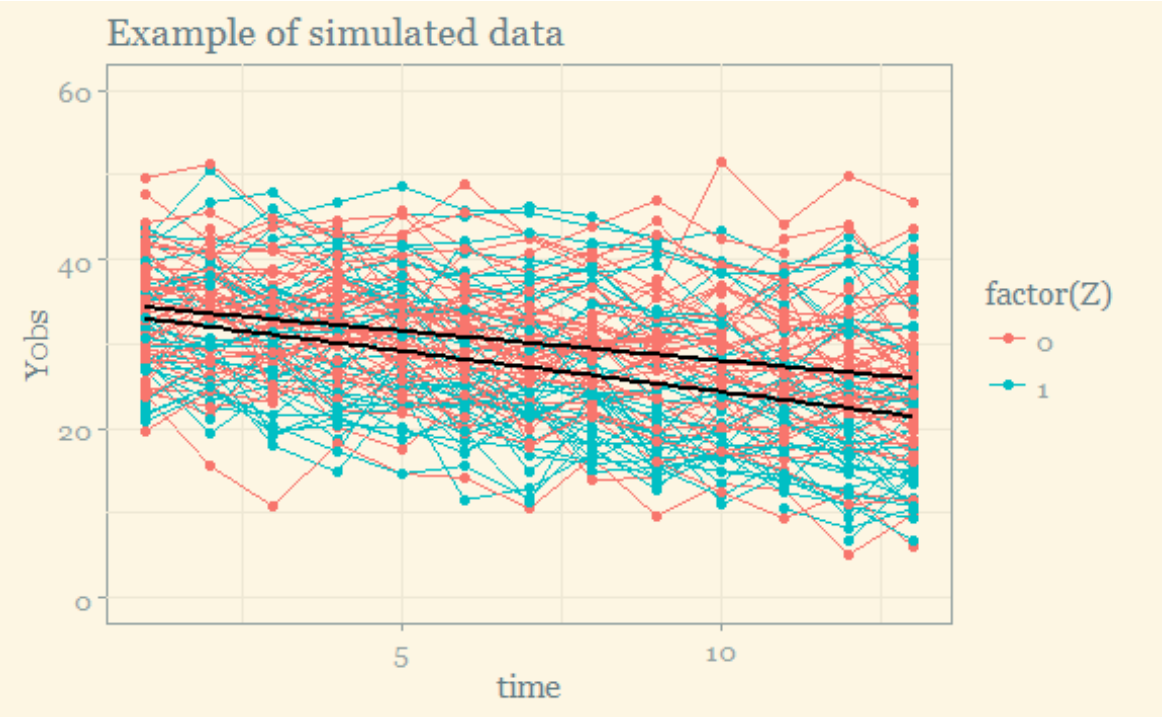
These parameters (interaction terms) can be difficult to interpret, so it is also good to present the effects within each treatment group to determine which one is more effective. The next plot shows the average change in HAI score over 12 weeks by treatment group, where the average change is estimated using the different methods as above. It appears that the KBT treatment group, on average, had a greater decline in HAI score over 12 weeks.



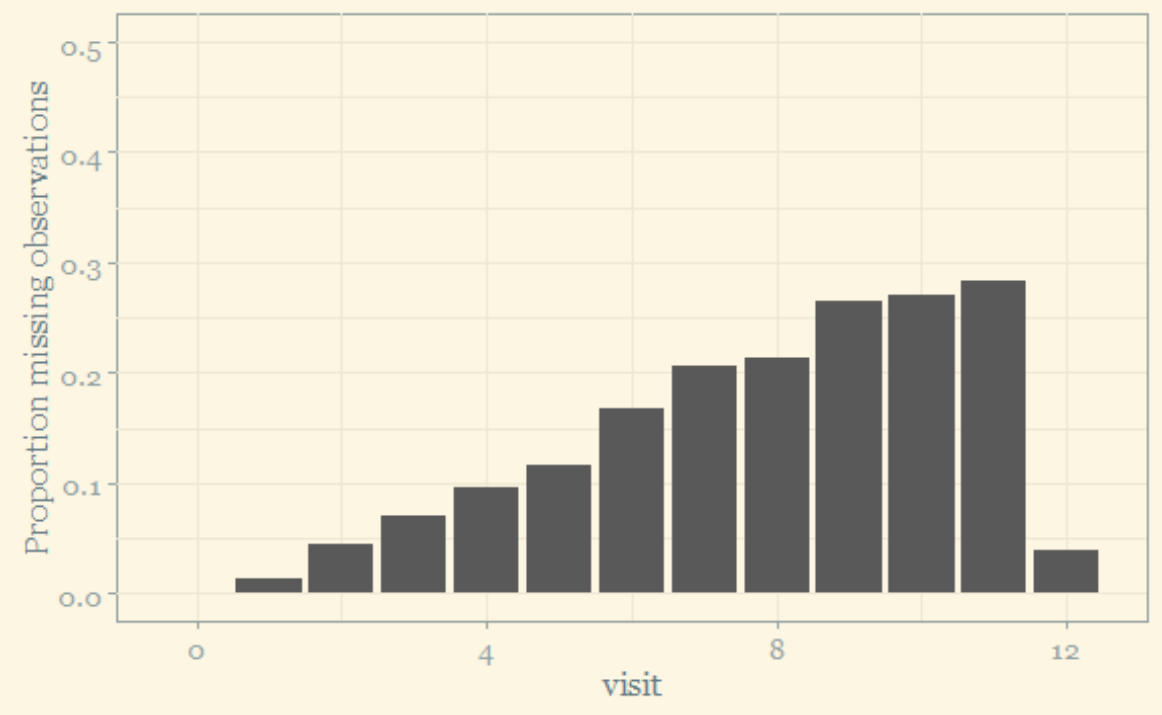
Power and Sample Size

A simulation based approach was used to evaluate power and sample size. The aforementioned data was used as guide on the data-generation mechanism. Briefly, the linear mixed effects model described in the equation above was used to generate hypothetical observations, using the estimated values from the previous study as parameter values. The data were generated using different values of the sample size, the value of γ (our parameter of interest), and different missing data mechanisms.

For each set of parameters, we generated data for a hypothetical trial. It was analyzed using the mixed effects model as described above, with Wald based confidence intervals, and we looked to see if the upper limit of the confidence interval excluded the noninferiority margin. If it excluded the margin, then the trial was considered a success. This procedure was replicated 5000 times, and the proportion of successes gives us an estimate of the power under those conditions. The next plot shows data from a single replicate of the experiment. Compare it to the real data set shown above.



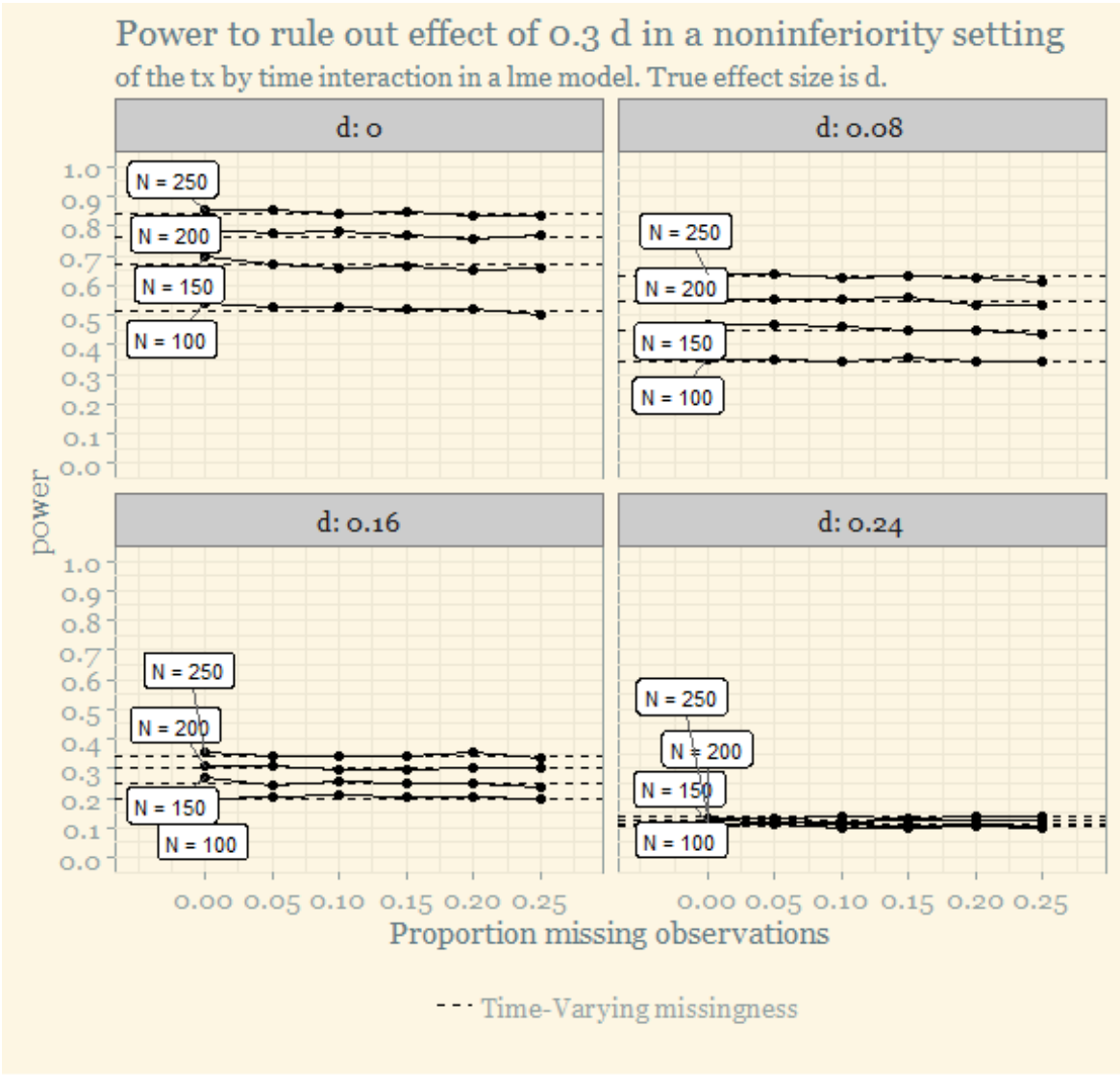
Several different missing data mechanisms were explored. First, different proportions of missing data were assumed to be uniform over the visit times. Second, we used the distribution of missingness from the prior study, which was clearly not uniform over the visit times (see below).

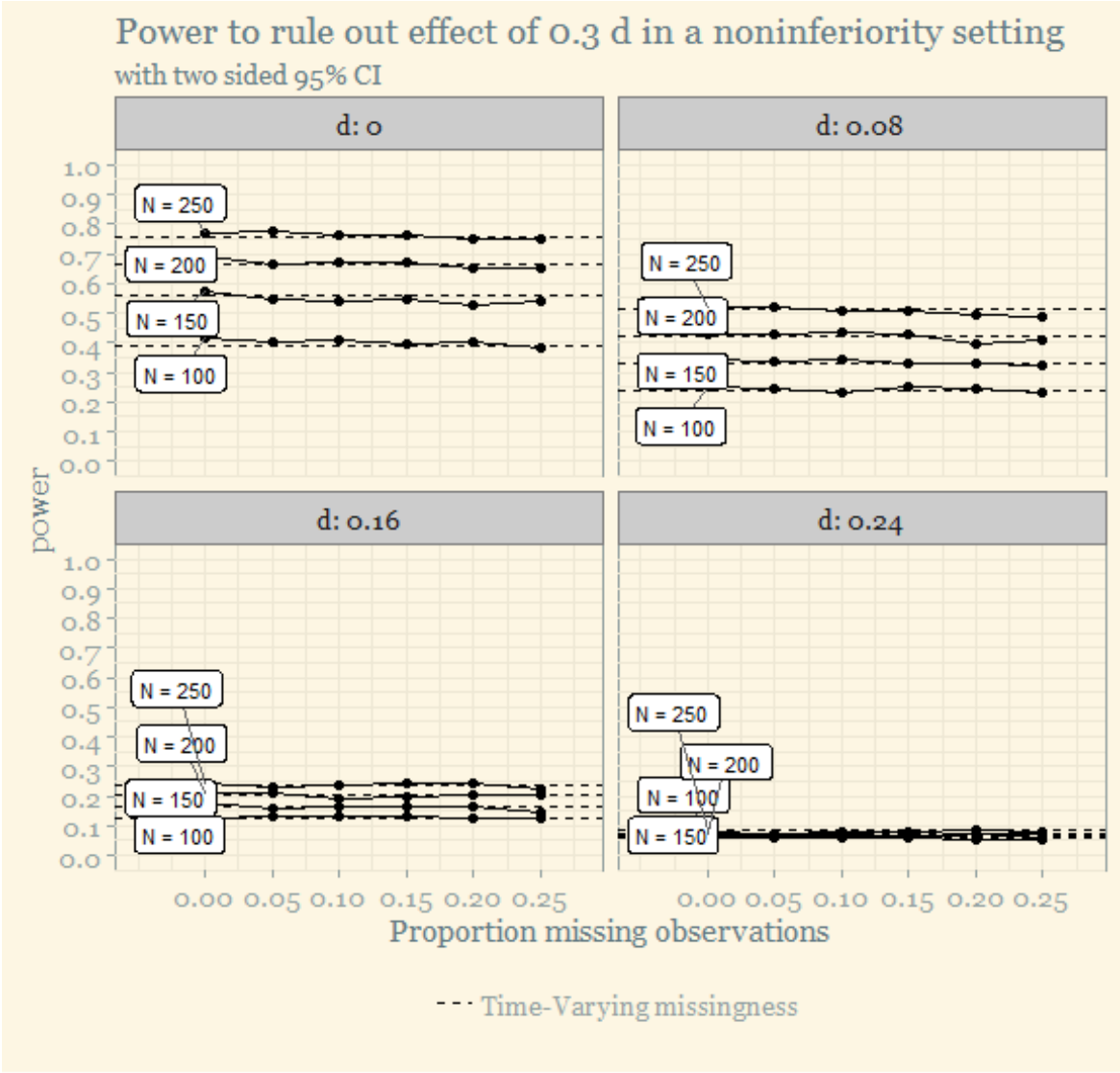


Results

Each line below represents a different sample size, ranging from 100 to 250. The 4 panels correspond to 4 values of the treatment difference of interest (the interaction term) on the Cohen's d scale. A larger value of d means that the experimental treatment is inferior to the standard treatment, but below the noninferiority margin. The dotted lines represent the scenario with missing data generated the same way it was observed in the previous trial.

For a sample size of 200 individuals total, and a true treatment difference of 0, there is approximately 80% power to confirm noninferiority, using a one-sided 95% confidence interval to rule out the margin. To achieve the same power with a true treatment difference of 0.08 would require greater than 250 individuals. As d approaches the noninferiority margin, the probability of declaring noninferiority approaches the type I error rate. Compared to a two-sided test, 200 individuals and a treatment difference of 0 only achieves 70% power.


Power to rule out effect of 0.3 d in a noninferiority setting
of the tx by time interaction in a lme model. True effect size is d.
d: 0
d: 0.08
d: 0.16
d: 0.24
N = 250
N = 200
N = 150
N = 100
power
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.00 0.05 0.10 0.15 0.20 0.25
Proportion missing observations
Time-Varying missingness



Appendix

Additional Resources (Links)

1. Reporting of Noninferiority Studies
2. Tutorial on linear mixed effects models, Part 1
3. Tutorial on linear mixed effects models, Part 2
4. Mixed models in SPSS

Code

```
library(rio)
library(tidyr)
library(stringr)
library(lme4)
library(dplyr)
```

```r
library(purrr)
library(ggplot2)
library(ggthemes)
library(ggrepel)
library(extrafont)
loadfonts(device = "win")

theme_set(theme_solarized(base_family = "Georgia"))

has <- import("../Data/Health anxiety example data to Michael with password.csv")

has$`HAI-M_shai_Veckom.[0]` <- has$`HAI-L_shai_Pre HAX`
has$`HAI-M_shai_Veckom.[12]` <- has$`HAI-L_shai_POST`
haslong <- has %>% gather(visitcode, hai_score, starts_with("HAI-M_shai_Veckom."))
haslong$visit <- as.numeric(str_match(haslong$visitcode, ".*\\[([0-9]+)\\]")[, 2])
ggplot(haslong, aes(x = visit, y = hai_score, color = (`TX Group`), group = `Internt ID`)) +
  geom_line(alpha = 0.6) + geom_point() + ggtitle("Distribution of HAI score over time") +
  stat_smooth(method = "loess", se = TRUE, aes(group = `TX Group`)) + scale_x_continuous(breaks = 0:12) +
  scale_y_continuous(limits = c(0, 60), breaks = seq(0, 60, by = 10))
sd.raw <- sd(has$`HAI-L_shai_POST`, na.rm = TRUE)
sd.diff <- sd(has$HAI_DIFF, na.rm = TRUE)

test.raw <- t.test(has$`HAI-L_shai_POST` ~ has$`TX Group`)

has$HAI_DIFF <- has$`HAI-L_shai_POST` - has$`HAI-L_shai_Pre HAX`
test.diff <- t.test(has$HAI_DIFF ~ has$`TX Group`)

hasslope <- haslong %>% group_by(`Internt ID`) %>%
  do({
    data.frame(haislope = lm(hai_score ~ visit, data = .)$coefficients[2] * 12,
               TX_group = .$`TX Group`[1], stringsAsFactors = FALSE)
  })

sd.slope <- sd(hasslope$haislope, na.rm = TRUE)

test.slope <- t.test(haislope ~ TX_group, data = hasslope)
```

```r
## mixed effects model
fitlme <- lmer(hai_score ~ visit * I(`TX Group` == "BSM") + (1 + visit | `Int
ernt ID`), data = haslong)
lme.est <- data.frame(point = fixef(fitlme)[4] * 12,
             lower = (fixef(fitlme)[4] * 12 - 1.96 * sqrt(diag(vcov(fitlme)))
[4] * (12)) ,
             upper = (fixef(fitlme)[4] * 12 + 1.96 * sqrt(diag(vcov(fitlme)))
[4] * (12)) )

pte <- function(test.obj) {

  data.frame(point = diff(rev(test.obj$estimate)),
             lower = test.obj$conf.int[1],
             upper = test.obj$conf.int[2])

}

resplo <- list(test.raw, test.diff, test.slope) %>% map_df(pte) %>%
  bind_rows(lme.est)

resplo$desc <- c("Only final visit", "Final - Baseline Difference", "Derived
Slopes", "Mixed Effects Model")

ggplot(resplo, aes(x = desc, y = point, ymin = lower, ymax = upper)) +
  geom_pointrange() + geom_hline(yintercept = c(0.0, 2.25), color = "red") +
  annotate("label", x = c(.7, .7), y = c(0.1, 2.25 + 1.25),
           label = c("No tx difference", "Noninferiority margin")) +
  scale_x_discrete(limits = resplo$desc) +
  xlab("") + ylab("Difference in 12 week change in HAI score") + ggtitle("Res
ults of different analysis methods", subtitle = "Point estimates and 95% conf
idence intervals") +
  coord_flip()

bytrt <- has %>% group_by(`TX Group`) %>%
  summarize(mnpost = mean(`HAI-L_shai_POST`, na.rm = TRUE) - mean(`HAI-L_shai
_Pre HAX`, na.rm = TRUE),
            sdpost = sqrt(var(`HAI-L_shai_POST`, na.rm = TRUE) + var(`HAI-L_s
hai_Pre HAX`, na.rm = TRUE) -
              2 * cor(`HAI-L_shai_POST`, `HAI-L_shai_Pre HAX`, use = "pairwis
e")),
            mndiff = mean(HAI_DIFF, na.rm = TRUE),
            sddiff = sd(HAI_DIFF, na.rm = TRUE))
```

```
res1 <- bytrt %>% gather("type", "mean", mnpost, mndiff) %>% gather("type2",
"sd", sdpost, sddiff) %>%
  mutate(typefin = substr(type, 3, 6), typefin2 = substr(type2, 3, 6)) %>%
  filter(typefin == typefin2) %>% select(TX_group =`TX Group`, mean = mean, s
d = sd, type = typefin) %>%
  mutate(lower = mean - 1.96 * sd / sqrt(nrow(has)), upper = mean + 1.96 * sd
/ sqrt(nrow(has)))

res2 <- hasslope %>% group_by(TX_group) %>%
  summarize(mean = mean(haislope, na.rm = TRUE),
            sd = sd(haislope, na.rm = TRUE),
            type = "slope", lower = mean - 1.96 * sd / sqrt(nrow(has)),
            upper = mean + 1.96 * sd / sqrt(nrow(has)))


bet <- fixef(fitlme)
cov <- vcov(fitlme)

res3 <- data.frame(TX_group = c("BSM", "KBT"), mean = c(((c(0, 1, 0, 1) %*% b
et) * 12)[1, 1],
  ((c(0, 1, 0, 0) %*% bet) * 12)[1, 1]),
  sd = c(12 * sqrt(c(0, 1, 0, 1) %*% cov %*% c(0, 1, 0, 1))[1, 1],
         12 * sqrt(c(0, 0, 0, 1) %*% cov %*% c(0, 0, 0, 1))[1, 1]),
  type = c("mixef", "mixef"))

res3 <- res3 %>% mutate(lower = mean - 1.96 * sd, upper = mean + 1.96 * sd)

allres <- bind_rows(res1, res2, res3)

desc <- c(post = "Only final visit",
          diff = "Final - Baseline Difference", slope = "Derived Slopes",
          mixef = "Mixed Effects Model")

ggplot(allres, aes(x = type, y = mean, ymin = lower, ymax = upper, color = TX
_group)) +
  geom_pointrange(position = position_dodge(width = .5)) +
  xlab("") + scale_x_discrete(limits = c("post", "diff", "slope", "mixef"),
                              labels = desc[c("post", "diff", "slope", "mixef
")]) +
  ylab("Average Change in HAI score over 12 weeks") +
  ggtitle("Results of different analysis methods",
          subtitle = "Point estimates and 95% confidence intervals") +
```

```r
  coord_flip()
load("example-plot.RData")
pex + ggtitle("Example of simulated data")
haslong %>% group_by(visit) %>% summarize(propmiss = mean(is.na(hai_score)))
%>%
  ggplot(aes(x = visit, y = propmiss)) + geom_bar(stat = "identity") +
  scale_y_continuous("Proportion missing observations", limits = c(0, .5))
load("sim-results-onesided-2016-08-30.RData")



res.power$d <- res.power$gamma * 12 / 7.5
labs <- res.power %>% group_by(gsize, d) %>% summarize(labyy = max(power), la
bxx = 0)
res.power$single <- sapply(res.power$propmiss, length) == 1

ggplot(subset(res.power, single), aes(x = unlist(propmiss), group = factor(gs
ize), y = power)) +
  geom_hline(data = subset(res.power, !single), aes(yintercept = power, linet
ype = "Time-Varying missingness")) +
  geom_line() + geom_point()  +
  geom_label_repel(data = labs, aes(x = labxx, y = labyy, label = paste("N ="
, gsize), group = NULL),
                   size = 3, nudge_x = -.1) +
  labs(title = "Power to rule out effect of 0.3 d in a noninferiority setting
",
          subtitle = "of the tx by time interaction in a lme model. True effe
ct size is d.") +
  scale_y_continuous(limits = c(0, 1), breaks = seq(0, 1, by = .1)) +
  scale_x_continuous("Proportion missing observations", limits = c(-0.05, .28
),
                     breaks = seq(0, 0.25, by = 0.05)) +
  scale_linetype_manual(values = c(2), guide = guide_legend(title = NULL)) +
  theme(legend.position = "bottom") + facet_wrap(~ d, labeller = "label_both"
)

load("sim-results-final-2016-08-27.RData")



res.power$d <- res.power$gamma * 12 / 7.5
labs <- res.power %>% group_by(gsize, d) %>% summarize(labyy = max(power), la
bxx = 0)
res.power$single <- sapply(res.power$propmiss, length) == 1
```

```
ggplot(subset(res.power, single), aes(x = unlist(propmiss), group = factor(gs
ize), y = power)) +
  geom_hline(data = subset(res.power, !single), aes(yintercept = power, linet
ype = "Time-Varying missingness")) +
  geom_line() + geom_point()  +
  geom_label_repel(data = labs, aes(x = labxx, y = labyy, label = paste("N ="
, gsize), group = NULL),
                   size = 3, nudge_x = -.1) +
  labs(title = "Power to rule out effect of 0.3 d in a noninferiority setting
",
          subtitle = "with two sided 95% CI") +
  scale_y_continuous(limits = c(0, 1), breaks = seq(0, 1, by = .1)) +
  scale_x_continuous("Proportion missing observations", limits = c(-0.05, .28
),
                     breaks = seq(0, 0.25, by = 0.05)) +
  scale_linetype_manual(values = c(2), guide = guide_legend(title = NULL)) +
  theme(legend.position = "bottom") + facet_wrap(~ d, labeller = "label_both"
)
```

Simulation Study

```
##

library(dplyr)
library(ggplot2)
library(ggthemes)
library(ggrepel)
library(lme4)
library(purrr)



gen_person <- function(neach = 13, alp = 34, gam0 = 0, bet = -1.0, gamma = 0.
0) {

  ttt <- seq(1, neach)
  Z <- rbinom(1, 1, .5)
  aaa <- rnorm(1, mean = 0, sd = 6)
  bbb <- rnorm(1, sd = .5)

  Y <- alp + aaa + gam0 * Z +
    (bet + bbb) * ttt +
    (gamma) * Z * ttt + rnorm(neach, sd = 3)

  data.frame(time = ttt, Z = Z, Y = Y)
```

```
}

gen_group <- function(gsize = 100, neach = 13,
                      alp = 33, gam0 = 0, bet = -.75, gamma = 0.0,
                      propmiss = 0.05) {   ## propmiss

  dat0 <- data.frame(pid = 1:gsize) %>% group_by(pid) %>%
    do(gen_person(neach, alp, gam0, bet, gamma))
  dat0 %>% group_by(pid) %>% mutate(miss = rbinom(neach, 1, propmiss),
                                    Yobs = ifelse(miss == 1, NA, Y))

}

panalyze <- function(dat0, marg = 2.25 / 12, outcome = "Yobs") {

  f1 <- " ~ time * Z + (1 + time | pid)"

  fit <- lmer(as.formula(paste(outcome, f1)), data = dat0, REML = FALSE,
              start = c(6, .5, 3), control = lmerControl(calc.derivs = FALSE)
)
  ci <- confint(fit, parm = 8, method = "Wald", level = 0.90)
  marg > ci[1, 2]

}


simulate <- function(B = 1000, param) {

  p0 <- sapply(param, unlist)
  replicate(B, {
    dat0 <- do.call(gen_group, as.list(p0))
    panalyze(dat0)
  }) %>% mean

}

set.seed(410)
dat0 <- gen_group(gam0 = 0)
system.time(panalyze(dat0))
```

```
pex <- ggplot(dat0, aes(x = time, y = Yobs, color = factor(Z), group = pid))
+ geom_line() + geom_point() +
  stat_smooth(method = 'lm', se = FALSE, aes(group = factor(Z)), color = 'bla
ck') + ylim(c(0, 60))

#save(pex, file = "example-plot.RData")

params <- cross_d(list(propmiss = c(0, 0.05, 0.1, 0.15, 0.2, 0.25),
                       gsize = c(100, 150, 200, 250),
                       gamma = c(0, 0.05, 0.1, 0.15)
                       ))

params$propmiss <- as.list(params$propmiss)

params.timevary <- data_frame(propmiss = lapply(1:16, function(i) {
  c(0, 0.01, 0.05, 0.07, 0.1, 0.12, 0.17,
    0.21, 0.21, 0.26, 0.27, 0.28, 0.04)}),
  gsize = sort(rep(c(100, 150, 200, 250), 4)),
  gamma = rep(c(0, 0.05, 0.1, 0.15), 4))

params.all <- bind_rows(params, params.timevary)

res.power <- params.all %>% by_row(simulate, B = 5000, .to = "power", .collat
e = "rows")

save(res.power, file = paste0("sim-results-onesided-", Sys.Date(), ".RData"))
```

Reproducibility Note

```
## R version 3.3.1 (2016-06-21)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 7 x64 (build 7601) Service Pack 1
##
## locale:
## [1] LC_COLLATE=Swedish_Sweden.1252  LC_CTYPE=Swedish_Sweden.1252
## [3] LC_MONETARY=Swedish_Sweden.1252 LC_NUMERIC=C
## [5] LC_TIME=Swedish_Sweden.1252
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] extrafont_0.17     ggrepel_0.5        ggthemes_3.2.0
##  [4] ggplot2_2.1.0.9000 purrr_0.2.2        dplyr_0.5.0
```

```
##  [7] lme4_1.1-12        Matrix_1.2-6       stringr_1.0.0
## [10] tidyr_0.6.0        rio_0.4.12         knitr_1.14
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.6      plyr_1.8.4       cellranger_1.1.0 formatR_1.4
##  [5] nloptr_1.0.4     tools_3.3.1      digest_0.6.10    gtable_0.2.0
##  [9] jsonlite_1.0     evaluate_0.9     tibble_1.1       nlme_3.1-128
## [13] lattice_0.20-33  openxlsx_3.0.0   csvy_0.1.3       DBI_0.4-1
## [17] curl_1.1         yaml_2.1.13      haven_0.2.1      Rttf2pt1_1.3.4
## [21] xml2_1.0.0       readODS_1.6.2    triebeard_0.3.0  grid_3.3.1
## [25] data.table_1.9.6 R6_2.1.2         readxl_0.1.1     foreign_0.8-66
## [29] rmarkdown_1.0    minqa_1.2.4      extrafontdb_1.0  readr_1.0.0
## [33] magrittr_1.5     scales_0.4.0     urltools_1.5.0   htmltools_0.3.5
## [37] splines_3.3.1    MASS_7.3-45      assertthat_0.1   colorspace_1.2-6
## [41] labeling_0.3     stringi_1.1.1    lazyeval_0.2.0   munsell_0.4.3
## [45] chron_2.3-47
```