International Journal of Haptonomy & Haptotherapy

Published in the International Journal of Haptonomy and Haptotherapy (08/10/17)

Haptotherapy and science

Revealing the secret of good research designs and valid measurement methods

Han CG Kemper

PhD, Emeritus Professor in Health and Activity, EMGO: Institute for Health and Care Research, VU University Medical Center Amsterdam, the Netherlands

Corresponding address:

Han C.G. Kemper, Meikevermeent 2, 1218 HD Hilversum, phone +31 35 6920710

E-mail address: hancgkemper@upcmail.nl

Abstract

Haptotherapy is a relatively new therapy with little sound scientific research to substantiate its effectiveness on health and well being of the clients. Unbiased and well-controlled research is sorely needed. In this article, a few research-related topics will be discussed with regard to haptotherapeutic treatment. The five topics to be discussed, include, measurement methods, research designs, quality assessment of haptotherapeutic research, (mis-)use of statistical analysis and meta-analysis.

Measurements used to evaluate health effects from haptotherapy must be reproducible and objective. These are prerequisites for reliable measurements. But the most important condition is that the measurement is valid, i.e. it measures what it is meant to measure.

Several research designs are available to investigate the effects of haptotherapy. Quasi-experimental designs can be used if randomization is not possible as is the case in most field experiments. The most adequate experimental design is the Randomized Controlled Trial (RCT) with randomized intervention and control groups. Thus far, only one RCT has been conducted in the near past (Van den Berg et al, 2006).

In the final two sections, the use of correlations and the advantages of meta-analysis are discussed.

1. Introduction

In daily practice of health care, an enormous variety of treatments are offered. For many treatments guidelines have been developed based on their effectiveness and possible side effects. This so-called evidence-based medicine or at least accepted methods are applied by workers in health care in the Netherlands who are registered under the Law on Professions in Health Care for the BIG (Wet op de Beroepen In de Gezondheidszorg) - such as medical doctors, medical specialists, physiotherapists, dentists, health care psychologists, psychiatrists and nurses.

Until now caregivers in alternative medicine, such as homeopathy, acupuncture, osteopathy and orthomolecular medicine, have not been forced to prove their methods as in regular medicine. Haptotherapy is still a therapy that is not accepted as evidence-based (Sackett et al, 1996) therapy in health care in the Netherlands because of a lack of research that indicates its effectiveness.

The major purpose of this article is to outline a number of research-related topics that need to be addressed in the development of a sound database in order to answer questions about the causal effects of their therapy. We will refer to these causes as treatments or independent variables and to the effects of the treatment as dependent variables or outcomes (Campbell and Stanley, 1963). Confounding variables can be interacting between the independent and dependent variables as shown in figure 1.

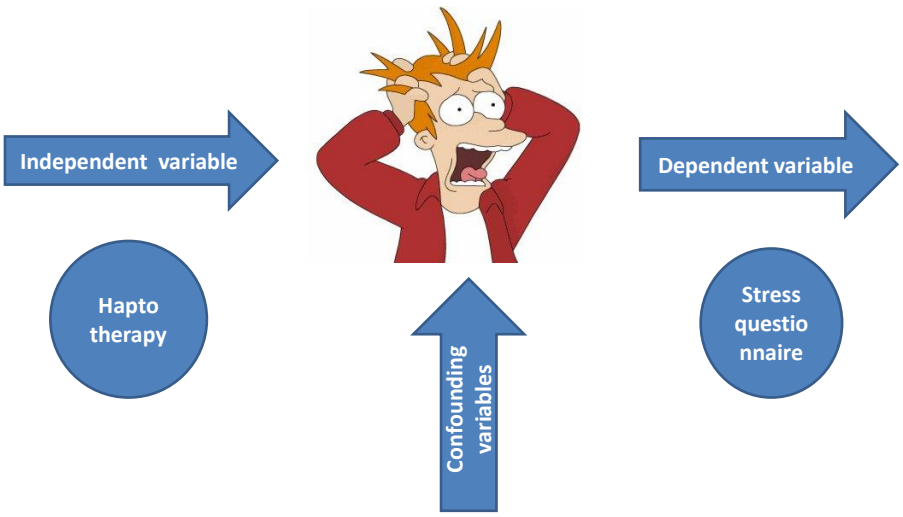

Figure 1: Principles of experimental research with independent, dependent and confounding variables: example of haptotherapy to reduce stress measured with a questionnaire

2. Reliable measurement methods

Good research starts with reliable measurements. If one is interested in the effects of haptotherapy in patients with stress, than stress has to be operationalized and measured by instruments that can be trusted such as questionnaires, measurement of hormones or sweat secretion.

But each of the methods has its pro's and cons: some are simple and cheap such as questionnaires. Others are invasive or not pleasant for the subjects involved because they need to use blood samples. The most important issue of measurements is that they are reliable. This concept of reliability of a measurement can be illustrated by a metaphor. A reliable measurement can be seen as an arrow that hits the target. The target can be seen as the concept we want to measure in the patient. The ultimate goal is that all the arrows finish

in or close to the target. If all the arrows stick into the target, the thrower is very precise and the method is valid, and therefore the measurement is measuring what it is supposed to measure (figure 2).

A valid measurement can only be reliable when it is also reproducible (intra tester reliability) and if the measurement is executed by different testers, it must also have a high intertester reliability). In figure 2 four situations with low and high reliability measures are illustrated. In the upper left panel (A) a valid measurement is shown, (always hitting the target) with high reproducibility. In the lower left panel (C) the measurement is not valid (no hits) but at the same time there is a high reproducibility. In the upper right panel (B) the measurement is valid (most of the arrows are in or close to the target) but not very reproducible. In the lower right panel (D) a measurement is shown that is neither reproducible (wide range of no-hits) nor valid, because the target is seldom hit.

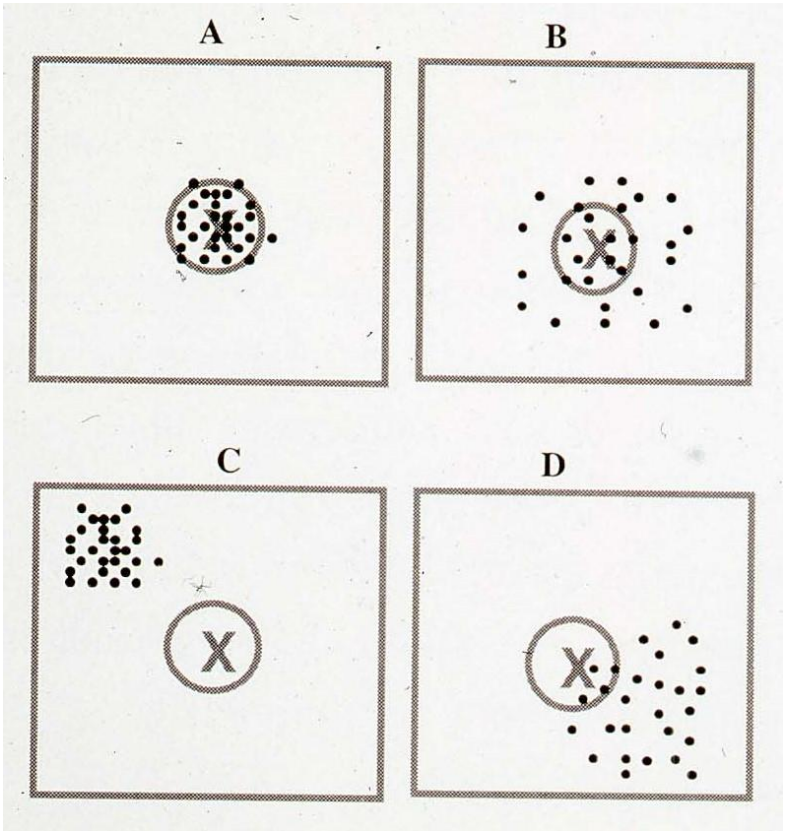

Figure 2: The reliability of a measurement method is illustrated as a metaphor with arrow shooters that try to hit the target (X)

In an experiment always all measurement methods have to be checked in advance or in earlier research on their reliability taking into consideration the population that is involved. Questionnaires are valid in adult subjects, but maybe less in males or females and not applicable in young subjects (Montoye et al, 1996). Stress hormone levels are different in children and adolescents compared with adults; the same holds for sweat secretion, which is dependent on body segment and age (Kemper, 2012)

3. Adequate research designs

The design of a research is the second important aspect to perform and publish successful scientific and proof research. It outlines the experimental approach to perform research about the causal effects of haptotherapy.

This paragraph discusses three pre-experimental and three true-experimental designs that are often used the practice of (hapto) therapy. The three pre-experimental designs attempt to partition respondents into non-equivalent groups that receive different treatments or no explicit treatment at all. We follow the notational system of Campbell and Stanley (1969) in which **X** stands for the exposure of a group to an experimental variable or event, in this case it can be a haptotherapeutic treatment. **O** will refer to some process of observation or measurement for an observation, f.i. a psychosocial questionnaire and/or a physical measurement such as sweat rate or cortisol level in blood. The left to right dimension indicates the temporal order of **X**'s and **O**'s. A prefix attached to observation **O** indicates a pretest ($\mathbf{O_{1)}}$ and a posttest ($\mathbf{O_2}$) in the annotation of the designs. Vertical symbols to one another are simultaneous.

The first three pre-experimental designs are frequently used in social sciences research. They are often used because they are simple and easy to comply, but they are normally not sufficient for permitting strong tests of causal hypothesis. The main reason is that there is no random assignment (symbol **R**) to the separate treatment groups and therefore they fail to rule out a number of plausible alternative interpretations (Cook and Campbell, chapter 3, 1979).

3.1 The One-Group Posttest Only Design

This design, also named the one-shot case study, involves making observations (O) only on persons who have undergone a therapeutic treatment (X), and then only after they have received it. This design is totally informative: there is a lack of pretest observations and an absence of a control group who did not receive the treatment.

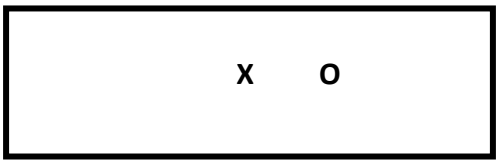

Diagram 1: The one-group posttest-only design

3.2 The Posttest-Only Design with Non-equivalent Groups

In therapy it often happens that a treatment is already implemented before the researcher can prepare for it, and so the research design is worked out after the therapeutic treatment has started (so called ex post facto). The most obvious flaw is the absence of a pretest. This leads to the possibility that any posttest difference between the two groups can be attributed either to the treatment effect or to selection or recruitment differences between groups. The plausibility of selection differences in research with non-equivalent groups usually renders this design uninterruptable.

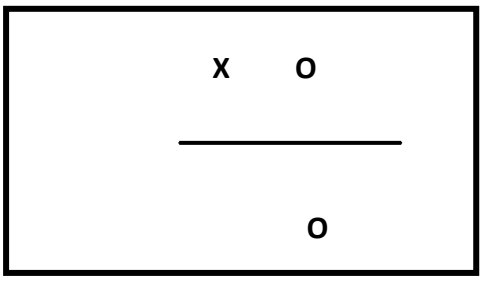

Diagram 2: The posttest-only design with non-equivalent groups

It is obvious that both designs (3.1 and 3.2) are not adequate to prove causal effects of their haptotherapy practice or intervention.

3.3 The One-Group Pretest-Posttest Design

This design is one of the more frequently used designs in social sciences. It can be seen that pretest observations (**O_1**) are recorded on a single group who later receive a therapeutic treatment (**X**), after which posttest observations are made (**O_2**). The differences found between pre- and posttest can also be caused by time dependent changes and not necessarily by the therapeutic intervention. One example is back pain: back pain is very common in our society and many therapeutic and medical treatments are used to cure this chronic disease. However it is also demonstrated that in many cases the back pain problems also disappear without any medication or physiotherapy. This is well known and underlined by the advice: if you have back pain, go back to work!

Care must also be taken if the expected posttest results are different from the pretest scores. This might be either because the posttest was affected by the treatment or because knowledge or motivation (not related to the **haptotherapeutic** intervention) gained at the first testing altered the subsequent testing results. This is the result of a confounding variable (see figure 1).

Therefore also this design is unable to prove that changes between **O_1** and **O_2** are the result of the haptotherapeutic intervention.

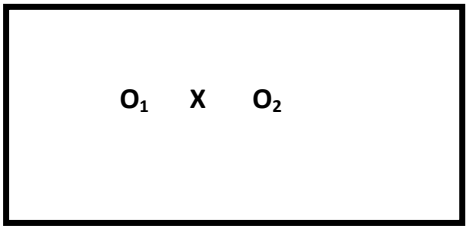

Diagram 3: The One-group Pretest-Posttest Design

The next three designs are called true experimental designs because they are using groups of subjects that are selected at random before the start of the experiment. R indicates this in the research diagrams at the left side.

3.4 The Randomized Pretest and Posttest

This design (also known as Randomized Controlled Trial, RTC) is the most used design in (para)medical and social sciences, because it is very interpretable and can therefore also recommended in (hapto) therapeutic situations. No differences between the treatment and control group at pretest can be considered if the two groups are selected at random (R). However this selection is sometimes in practice difficult to achieve. If the selected groups are big the probability of identical groups is higher than in small groups.

Also comparison of school classes that are treated or not can give problems because within the groups different interactions can cause different changes, not related to the therapeutic intervention.

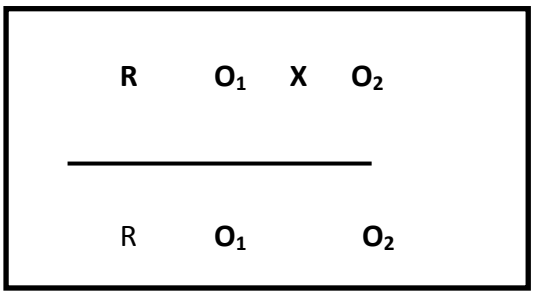

Diagram 4: The Randomized Pretest and Posttest Control Group design

3.5 The Randomized Four Groups Design with or without Pretest and Posttest

This design can reveal if the pretest in itself can have any influence on the posttest results as was hypothesized in the earlier research designs. By paralleling the design 3.4. Elements (**O_1**

through O_4) with experimental and control groups lacking the pretest (O_5 and O_6), both the main effect of pretesting and the interaction of pretesting and **X** are determinable.

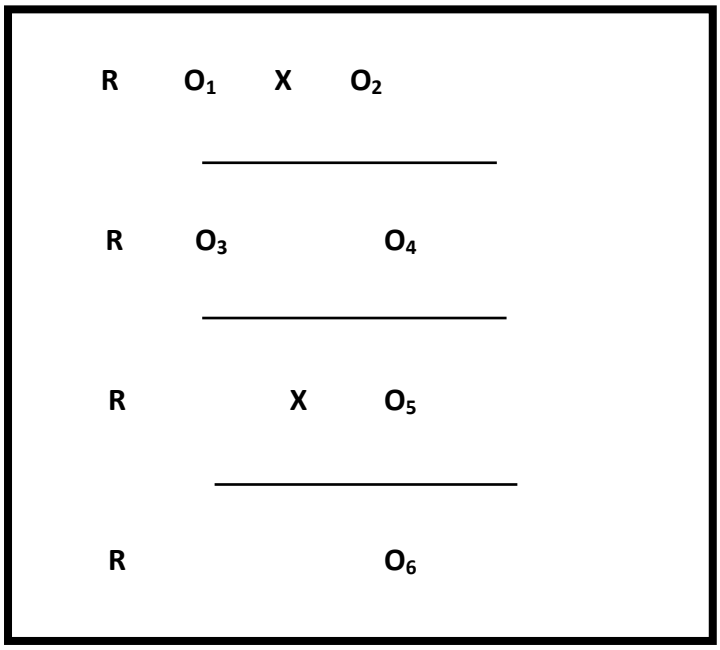

Diagram 5: the Randomized Four Groups Design with or without Pretest and Posttest

3.6 The Randomized Two Groups Design without Pretest

This last design can be used if can be supposed that the random chosen groups are not different on the important dependent variables. Intervention in children can be influenced by the individual rate of growth and maturation of the subjects.

In the daily practise it is more difficult to assign individuals to treatments or therapies at random in field settings than in laboratory settings.

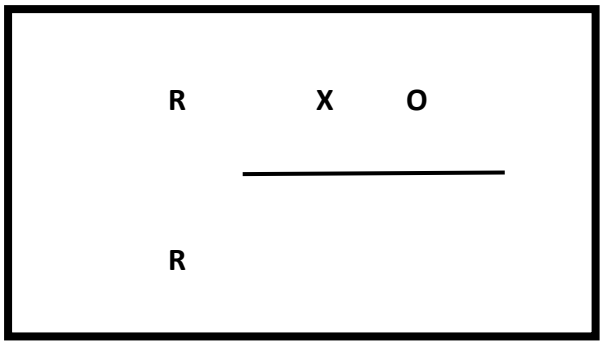

Diagram 6: the Randomized Two Groups Design without Pretest

The three true experimental research designs described here are only the most used or most practical in field settings.

More designs are described elsewhere (Cook and Campbell, 1979). They include (a) designs where experimental and control groups are alternately exposed to the same treatment **X**; the control group is first used as a waiting group and the treatment is given to this control group after finishing the first experiment. (b) In some therapeutic research the design allows for a reversed treatment i.e. a treatment **X+** that is expected to have an effect in one direction and a treatment **X-** that has an effect in the opposite direction. (c) A repeated treatment design in which in one population the treatment is introduced and faded out and then reintroduced at a later date.

4. Quality assessment of haptotherapeutic research in the near past

Included in this paragraph are six (intervention) studies concerning haptotherapeutic therapies. This collection of published studies shows that hardly any interventions are published. One of them is without intervention and most of the others are not published in refereed journals. The studies are published between 1997 and 2012. A seventh multi centre trial by Klabbers et al (2014) about treatment of severe fear of childbirth with haptotherapy is still in progress.

In order of publication the six studies in short: In 1997 Van 't Hof et al published "Denken over Voelen". There was no pretest neither a control group. Van Luttervelt published in 1997 a philosophical literature survey "Bevestigend Aanraken". It was no effect study. Akkerman published in 2005 "Haptotherapie en copingstijl". It showed in neck/shoulder patients a positive effect, but there was no control group involved.

Van der Berg et al published in Journal of Patient Education and Counseling (2006) an intervention "Evaluation of Haptotherapy" in cancer patients undergoing chemotherapy. They used a valid design the randomized pre- and posttests control group design (see diagram 4) and their results showed weak intervention effects. De Frangh wrote a master thesis in Leuven (2009) called "Empatische Toets van Eigenheid en Haptotherapie" with self-reports of clients and their therapists. There design involved only a posttest (see diagram 1) and no effects are found. Bosscher et al did a study "Clinical Effectiveness of Haptotherapy" in 2013. This intervention study used a one-group pretest posttest design (diagram 3) but their intervention effects are small.

5. (Mis-)use of statistical analyses on the collected data

In evidence based medicine the results are analysed by statistical tests. The aim is to show that a treatment works or not. It is first hypothesized that there is no effect (Thero hypothesis, H_0). In medicine most of the time a chance (p) of 5 or 1 per cent is chosen to decide that there is an effect. This is called the level of significance ($p < 0.05$ or $p < 0.01$).

Apart from a statistical significant effect it is also important to know how big this effect is. A highly statistical significant effect can be so small that it is not a relevant therapeutic effect. If a therapy significantly decreases the average stress (on a scale from 1 to 20) from 12 to 11 we may call this as a clinically not significant.

A correlation coefficient is a number between -1 and +1(fig 3). A correlation around 0 means no relation at all between the two variables (Y and X) involved. A plus one correlation ($r = +1$) indicates a perfect positive relationship (the longer the height of haptotherapists the bigger their hands or feet) and a minus one correlation ($r = -1$) a negative relationship between the

two variables (the higher IQ of students the less time they use for their homework). A 0.5 correlation means that the relation between Y and X can explain 25% of the variation.

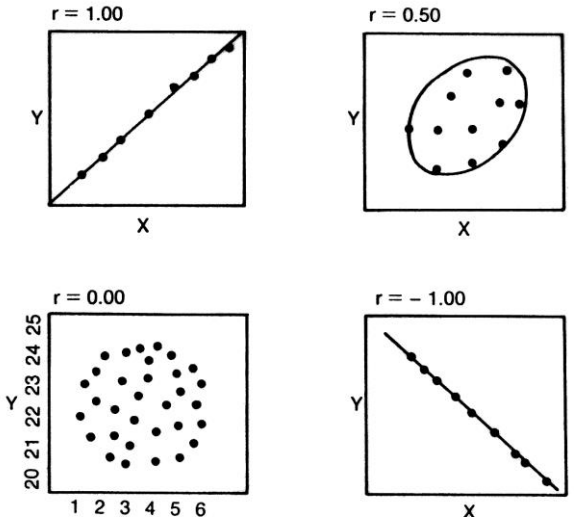

Figure 3: Correlations (r) between Y and X can vary from plus one (upper left) to minus one (lower right)

The use of correlations is another misuse in statistics. Often the correlation measure between two variables is interpreted as a causal relationship. If in a research population two variables are measured, such as stress and workload, a high correlation between them cannot be interpreted as “a high workload causes a high level of stress”. The direction of this relation is not known, because it is also possible that it is the other way around: ”stressful subjects experience their work as high ”.

Moreover high correlations between two variables can also be caused by another third variable that is correlated with both other two variables. An example: in a big population of children between 6 and 18 years stress (experienced on a scale from zero to ten) is correlated with sweat rate. The relation is linear and highly significant.

However in children the number of sweat glands is increasing with age and the same holds for the experienced stress. The graph in figure 3 (Rowland, 1990) suggests a positive relation

between stress (vertical axis)) and sweat rate (horizontal axis). This is caused by a third variable the age of the children age. Both stress (vertical Y-axis) and sweat rate (horizontal X-axis) are higher in older children (figure 4).

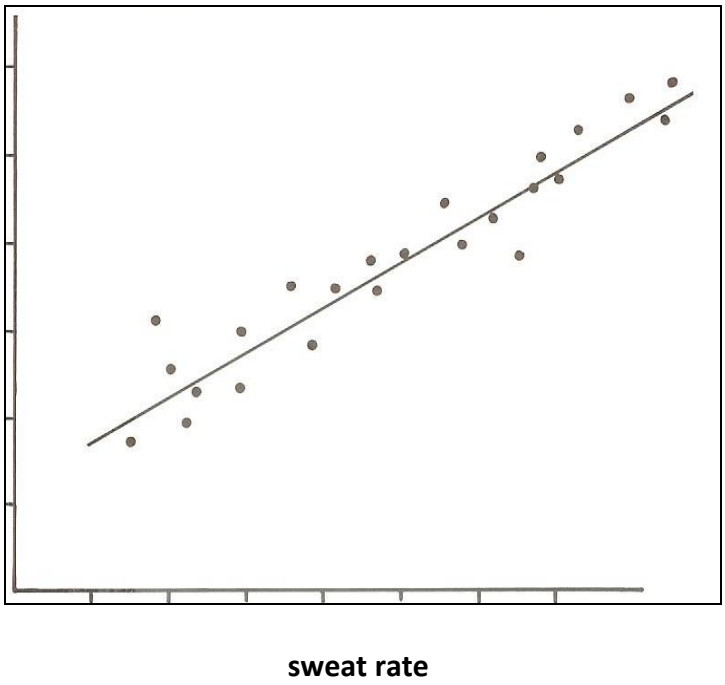

Figure 4: A high and significant correlation between the two variables tells nothing about cause and effect

Although a correlation can be significant, it cannot always be used in individual therapeutic situations. A correlation can become significant because it is calculated from a population that is very diverse in the variables measured (figure 5). In the left figure is shown that the linear relationship between aerobic power (VO2max) and the performance on a performance run is significant (0.53) and positive, but at the same time the individual variation (indicated by the vertical and horizontal arrows) around this calculated regression line is also big.

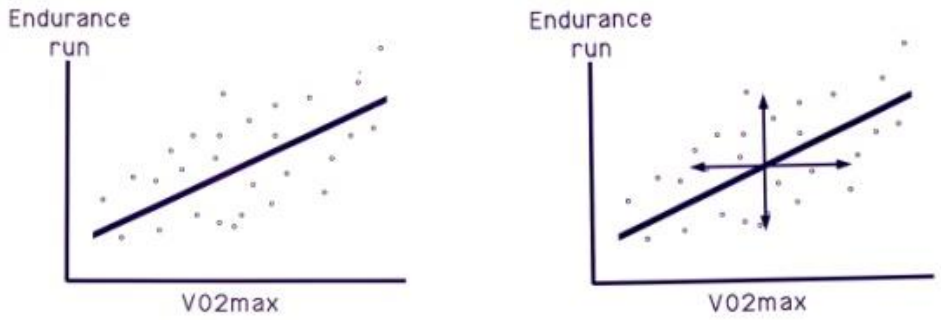

Figure 5: The overall significant linear correlation shows a big variation in individuals

6. Advantages of meta-analyses

Today where it is easy to gather by internet the results and data collected by different researchers there is an important trend to do e research synthesis of results from different studies on the same topic and wit the same research design (Cooper et al, 1994).

A research synthesis is also called a meta-analysis. A descriptive or narrative meta-analysis compares different studies, makes a quality assessment of each, selects the best ones and looks at the results. When there are many small or weak studies of an intervention, a statistical meta-analysis can be used to co-ordinate the study results and to draw a stronger conclusion about the outcome of the treatment. This can be an important contribution to the establishment of a foundation of evidence about an intervention.

The strongest meta-analysis is the systematic review in which the data from the included studies are integrated in one database and analysed. More equivalent data from many relatively small studies with insignificant effects have the advantage tot get significant results in the secondary analysis of the big database (Wolff et al, 1999).

Meta-analyses have also a disadvantage: research articles with non-significant results are difficult to publish because editors of refereed journals are more inclined to accept

manuscripts with significant results. In that way the non-significant studies are underrepresented in the database of meta-analytical studies.

In the case of haptotherapy we have seen that not much (valid) research is performed even in the Netherlands and Belgium, and not at all in other countries. Therefore a meta-analytical approach to reveal the therapeutic effects of haptotherapy is not possible.

7. Conclusions

In this article has been revealed the important steps to perform valid scientific research in understanding the therapeutic effects of haptotherapy. It all starts with developing reliable measurement methods and to apply these in adequate research designs. In the near past only a limited number of haptotherapeutic researches have been done and there is need for high quality research with scientific proof designs in order to get evidence-based results of the effects of haptotherapy.

My hope is that this article stimulates the haptotherapeutic world to set-up valid scientific intervention studies to reveal the true effects of their therapy. Developing a database of therapeutic effects can support a meta-analysis in the near future.

This article is based on an earlier lecture (Haptotherapie en Wetenschap: Wat zijn de haken en ogen?) by the author at the Congress Haptotherapie Nederland on Saturday March 16 in 2013 at the Free University in Amsterdam (NL).

8. References

Ackerman, H. (2005). *De invloed van haptotherapie op de copingstijl van werknemers met nek- en schouderklachten.* Open Universiteit Nederland, Faculteit Psychologie, Heerlen.

Berg, M., Visser, A., Schoolmeesters, A., Edelman, P., & Borne, P. (2006). Evaluation of haptotherapy for patients with cancer treated with chemotherapy at a day clinic. (PubMed, Ed.) *Patient Education and Counseling , 60*, 336-343.

BIG. (2015). Wet op Beroepen In de Gezondheidszorg. The Hague: Ministry of Health, Welfare and Sport.

Bosscher, R., Leeuwen, A., & Pluimers, C. (2013). The clinical effectiveness of haptotherapy in routine practices. *International Journal of Haptonomy & Haptotherapy*.

Campbell, D., & Stanley, J. (1967). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally and Company.

Cook, T., & Campbell, D. (1979). *Quasi-experimental, design and analysis issues for field settings.* Chicago: Rand McNally Publishing Company.

Cooper, H., Hedges, L., & Valentin, J. (1996). *The handbook of research synthesis and meta-analysis.* New York: Russel Sage Foudation.

Frangh, A. (2009). *Definiëring en empirische toetsing van eigenheid van haptotherapie als lichaamsgerichte therapievorm.* Leuven: Master thesis; Katholieke Universiteit.

Hof, A., Voorst, P., & Zoelen-Nederlof, M. (1997). *Denken over voelen, een onderzoek naar de kenmerken van de cliëntenpopulatie van de Vereniging Van Haptotherapeuten (VVH) en de betekenis van haptotherapie voor deze cliënten.* Wetenschapswinkel, Sociale Wetenschappen. Universiteit Utrecht.

Kemper, H. (2012). *Fitte kinderen, sportieve tieners. Over de invloed van bewegen op de gezondheid van jongeren.* Amsterdam: Elsevier Gezondheidszorg.

Klabbers, G., Wijma, K., Paarlberg , M., Emons, W., & Vingerhoets, A. (2014). Treatment of severe fear of childbirth with haptotherapy: design of a multicenter randomized controlled trial. *The offical journal of the International Society for Complementary Medicine Research (ISCMR) , 14:* (385) DOI**:** 10.1186/1472-6882-14-385.

Luttervelt, M. van (1998). *Bevestigend aanraken.* Master thesis, Erasmus Universiteit, Rotterdam.

Montoye, H., Kemper, H., Saris, W., & Washburn, R. (1996). *Measuring physical activity and energy expenditure. Human Kinetics.* Champaign (IL): Human Kinetics.

Rowland, T. (1990). *Exercise and Children's Health.* Champaign (IL): Human Kinetics Books.

Sackett, D., Rosenberg, J., Gray, J., Haynes, R., & Richardson, W. (1996). Evidence based medicine: what it is and what it is not. *British Medical Journal, 1996; 312* (7023): 71-72.

Wolff, I., Croonenburg, J., Kemper, H.C.G., Kostense, P., & Twisk, J. (1999). The effect of exercise training programs on bone mass: a meta-analysis of publisched controlled trials in pre- and postmenopausal women. *Osteoporosis International* 9 (1): 1-12.