



# Acupuncture diagnoses and treatment recommendations for infertile and fertile women

*A study of inter-rater reliability on diagnosis and treatment recommendation, and an examination of the distribution of acupuncture diagnoses in infertile and fertile women*

Oddveig Birkeflet



Institute of Health and Society,  
Faculty of Medicine,  
University of Oslo,  
Norway 2015

© **Oddveig Birkeflet, 2016**

*Series of dissertations submitted to the  
Faculty of Medicine, University of Oslo*

ISBN 978-82-8333-244-5

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Cover: Hanne Baadsgaard Utigard  
Printed in Norway: 07 Media AS – [www.07.no](http://www.07.no)

## Table of contents

Acknowledgments.....	3
Abbreviations.....	4
Summaries.....	5
Summary in English.....	5
Background.....	5
The objectives.....	5
Methods.....	5
Results.....	6
Conclusion.....	6
Summary in Norwegian.....	7
Bakgrunn.....	7
Formål.....	7
Metode.....	7
Resultat.....	8
Konklusjon.....	9
Introduction.....	10
Development from old theories to modern TCM acupuncture.....	11
Yin and Yang.....	11
Internal organs – Zhang-Fu Organs.....	13
The meridians.....	14
Vital Substances.....	15
Identifying imbalances in nourishment.....	15
Pathological factors.....	16
Pattern differentiation – fundamental principle of TCM diagnostics.....	17
Treatment principles.....	18
Infertility and TCM.....	19
TCM patterns related to infertility.....	19
Biomedical influence on the evolution of Chinese medicine.....	20
Pattern identification as a tool to integrate CM and western medicine.....	20
One pattern – many diseases, one disease – many patterns.....	21
Significance of the patterns.....	22
Aims of the present study.....	24
Study design, participants and methods.....	25

Study Participants .....	25
Collecting data and taking case histories.....	26
Case histories .....	27
Background of the acupuncturists.....	28
Ethical considerations.....	28
Data analysis .....	29
Statistical methods.....	32
Main results .....	33
Inter-rater reliability of TCM patterns (Papers I and II) .....	33
Impacts of demographic variables on the frequency of TCM patterns (Paper II) .....	34
Variations in use of symptoms and signs to diagnose TCM patterns (Paper II) .....	34
Treatment recommendations (Paper I) .....	35
TCM patterns in fertile and infertile women (Paper III) .....	36
Prevalence of acupuncture points for fertile and infertile (Paper III).....	37
Discussion.....	39
Low inter-rater agreement on TCM patterns .....	39
Factors that could cause variations in diagnosis .....	40
Impacts of TCM theories on diagnosis.....	44
TCM patterns in fertile and infertile women .....	45
Treatment recommendations (Papers I and III).....	46
Conclusion.....	50
Reference List.....	51

## Acknowledgments

Thanks to my supervisor Professor Nina Vøllestad for demanding more specifications, always requesting me to express myself clearer and more precisely. To statistician Professor Petter Laake for introducing me to the world of statistics.

Thanks to all the nine acupuncturists that were brave enough to let their diagnosis be compared. Without your participation I could not have completed this thesis. Thanks also to all the fertile and infertile women that shares their symptoms and sings with us.

Also my initial supervisor that sadly deceased, Professor Margarethe Lorensen, needs to be remembered. She introduced me to clinical nursing and science, showed me the openness and tolerance for all opinions, as long as you argued with factual knowledge.

Thanks to PhD and physiotherapist Gry Sagli for reviewing parts of my thesis and giving advice.

I am grateful to my special, lovely family team that always is involved in each other's work, and acted as my research team. Øystein Graf, my husband, thanks for all your patient help with excel. Odin Birkeflet, oldest son, I appreciate your valuable opinions, critical reflections and discussions concerning methods in scientific articles. My dear youngest son Johannes Birkeflet, thank you for your technical- and IT support, and help with the PowerPoint illustration and film production about this work.

Thanks to my mother and my departed father – and their parents, for giving me the manners to undertake hard work through hard times. As in the spirit of Daruma - do never give up even when the possibility is closed; seven times down - eight times up, became my inspiration.

The study has been self-financed, but I am grateful for economic support from the Norwegian Acupuncture Foundation and "Norge Kina Samarbeidet" directed by The National Research Center in Complementary and Alternative Medicine (NAFKAM).

## Abbreviations

There are different interpretations of the concepts used in Traditional Chinese Medicine (TCM), both in China and worldwide. This thesis was written in accordance with the definition from the World Health Organization's standards.

The organs and concepts referred to in Chinese Medicine are each written with an initial capital letter to emphasise their different meaning compared to biomedicine.

### Abbreviations

CI	confidence interval
CM	Chinese Medicine
EBM	evidence-based medicine
IVF	in vitro fertilisation
k	kappa statistic
Max k	maximum kappa
MB	medical bachelor's degree
NAC	Norwegian Acupuncture College
OR	odds ratio
RCT	randomised controlled trials
TCM	Traditional Chinese Medicine
WHO	World Health Organization

## Summaries

### Summary in English

#### Background

Acupuncture is used in treatment of a lot of conditions, for instance female infertility. Treatment is based on diagnoses (or patterns) according to Traditional Chinese Medicine (TCM) theory. According to Maciocia, eight TCM patterns are regarded typical for female infertility. However, there is a lack of empirical data comparing these patterns in fertile and infertile women. Furthermore, there are some studies indicating low reliability of TCM diagnosis. The causes of the variability are almost unexplored, but may be related to recognition of symptoms and signs, educational background or other demographic variables.

#### The objectives

The aim of this study was to examine the inter-rater reliability of TCM diagnoses, acupuncture points recommended by acupuncturists and to identify factors with impact on the diagnosis. Moreover, to compare the TCM patterns and acupuncture point recommendation in infertile and fertile women. Also, to examine the symptoms and sign used to set one pattern diagnosis and explore the influence of demographic variables on diagnosis.

#### Methods

30 infertile and 24 previously pregnant women were examined by two acupuncturists. The two acupuncturists were together in the consultation, and independently decided on TCM patterns and the prescription of acupuncture points. A sample of 25 case histories (15 fertile, 10 infertile) was made from a random selection of the 54 women. Eight acupuncturists independently diagnosed the written case histories. They also reported the symptoms and signs they used to diagnose. The diagnosed patterns, from all acupuncturists, were categorized into Excess-, Deficiency- and Merged patterns for analysis.

The kappa statistic was used to examine the inter-rater reliability of TCM pattern diagnosis and the acupuncture point recommendation. The association between the diagnosis and the presence of symptoms was examined for a single TCM diagnosis, Liver Qi Stagnation. The impact of demographic factors on the diagnosis was assessed by odds ratios.

To compare the TCM patterns and recommended acupuncture points in infertile and fertile women, the six women with secondary infertile were excluded. The odds ratio (OR) was used as the effect measure to examine the distribution of TCM patterns and acupuncture points among 24 infertile and 24 fertile women.

### **Results**

Large variation in the number of TCM pattern diagnoses and poor inter-rater reliability on TCM patterns were found. Some acupuncturists did not diagnose Blood Excess or Heart Deficiency at all, whereas others used these on more than 10 of the 25 case histories. Even for the most frequently used merged patterns there was poor agreement.

The demographic variables sex, duration of practice, and education and working hours had a significant effect on the frequency of TCM patterns set by eight acupuncturists. Having a background as a nurse or a physiotherapist was associated with setting a higher number of TCM patterns compared to those who completed basic medical courses only.

There was considerable intra- and inter-rater variation in the use of symptoms to set a diagnosis. Symptoms occurring frequently as well as infrequently were inconsequently used by the eight acupuncturists in setting the diagnose Liver Qi Stagnation.

For the choice of single acupuncture points poor agreement were found for the two acupuncturists. When points were merged according to their respective meridians, the agreement improved to perfect to moderate agreement.

Only three of the eight patterns assumed typical for infertile were more frequently diagnosed among infertile. Two of the eight patterns were more frequent among fertile. Three patterns were found in more than 92% of the women. Some patterns occurred more frequently among infertile whereas others among fertile. Acupuncture points Ki3, Liv3 and Sp6 were chosen for almost all women, whereas Cv4, Cv3 and St29 were chosen more frequently for infertile and Sp3 more frequently for fertile.

### **Conclusion**

The results showed extensive variation and poor inter-rater reliability in TCM diagnoses. The symptoms were used inconsistently to set a diagnosis. The low agreement on diagnoses indicates that acupuncturists follow individual pattern differentiation processes, even though they have a common educational background with respect to acupuncture. The



selection of acupuncture points seem to be closely related to the choice of TCM pattern diagnoses. Our data indicate that the presence of some and the absence of other TCM patterns are associated to infertility. These results are not consistent with the assumed associations between eight specific patterns and infertility, as maintained by TCM textbooks. The inconsistencies and variability in TCM pattern diagnosis may impede the individually tailored treatment, and is thus a challenge for clinical practice.

## **Summary in Norwegian**

### **Bakgrunn**

Akupunktur brukes i behandlingen av en rekke tilstander, for eksempel kvinnelig infertilitet. Behandlingen er basert på diagnoser/mønstre, satt etter kriterier fra Tradisjonell Kinesisk Medisin (TKM). I følge en mye brukt kilde, Maciocia, er det åtte TKM diagnoser/mønstre forbundet med kvinnelig infertilitet. Imidlertid er det mangel på empiriske data som sammenligner disse diagnosene/mønstrene hos fertile og infertile kvinner. Dessuten er det noen studier som indikerer at det er lav reliabilitet for TKM diagnoser. Årsaken til variabiliteten er nærmest uutforsket, men kan relateres til gjenkjenning/fortolkning av symptomer og tegn, utdanningsbakgrunn eller andre demografiske variabler.

### **Formål**

Formålet med denne studien var å undersøke inter-rater reliabiliteten av TKM diagnoser, akupunktørens anbefaling av akupunkturpunkter og å identifisere faktorer med innvirkning på diagnosene. Dessuten å sammenligne TKM diagnosene og akupunkturpunkt anbefalingene gitt til infertile og fertile kvinner. Likeså å undersøke hvilke symptomer og tegn som ble brukt til å sette en diagnose, og å undersøke effekten av demografiske variabler på diagnostiseringen.

### **Metode**

30 infertile og 24 tidligere gravide kvinner ble undersøkt av to akupunktører som var sammen under konsultasjonen. Akupunktørene satte uavhengig av hverandre TKM diagnoser og anbefalte akupunkturpunkter. Fra et tilfeldig utvalg av de 54 kvinnene, ble det laget 25 casehistorier (15 fertile, 10 infertile). De skrevne casehistoriene ble diagnostisert av åtte akupunktører uavhengig av hverandre. De rapporterte også hvilke symptomer og tegn de hadde brukt for å diagnostisere. For å analysere, ble diagnosene fra alle akupunktørene kategorisert i «Excess patterns», «Deficiency patterns» og «Merged patterns».

Inter-rater reliabiliteten av TKM diagnosene og akupunkturpunktanbefalingen ble undersøkt ved hjelp av kappa statistikk. Forbindelsen mellom diagnose og symptomer og tegn ble undersøkt for en diagnose, «Liver Qi Stagnation». Demografiske faktorerers effekt på diagnostiseringen ble undersøkt ved odds ratio.

For å sammenligne TKM diagnosene og de anbefalte akupunkturpunktene hos infertile og fertile kvinner, ble seks kvinner med sekundær infertilitet ekskludert. Odds ratio (OR) ble brukt som effekt mål for å undersøke fordelingen av TKM diagnoser og akupunktur punkter blant 24 infertile og 24 fertile kvinner.

### **Resultat**

Det var stor variasjon i antallet TKM diagnoser og svak inter-rater reliabilitet om TKM diagnosene. Noen akupunktører diagnostiserte ikke «Blood Excess» eller «Heart Deficiency» i det hele tatt, mens andre satte disse diagnosene på mer enn 10 av de 25 casehistoriene. Selv for de hyppigst brukte sammenslåtte diagnosene var det svak enighet.

Demografiske variabler som kjønn, praksistid, utdanning og arbeidstimer hadde signifikant effekt på antallet TKM diagnoser satt av de åtte akupunktørene. Akupunktører som også var utdannet sykepleier eller fysioterapeut satte flere TKM diagnoser sammenlignet med de som hadde grunnfag i medisin.

Det var betraktelig intra- og inter-rater variasjon i bruk av symptomer for å sette en diagnose. Både hyppig og sjelden forekommende symptomer ble brukt inkonsekvent av de åtte akupunktørene for å sette diagnosen «Liver Qi Stagnation».

For valg av enkelt akupunkturpunkter ble det funnet svak enighet mellom de to akupunktørene. Når punktene ble sammenslått i forhold til deres tilhørende meridian, ble enigheten forbedret til perfekt til moderat enighet for de mest brukte meridianene «Liver», «Kidney», «Stomach» og «Spleen».

Kun tre av de åtte diagnosene forbundet med infertilitet ble diagnostisert hyppigst blant de infertile. To av de åtte diagnosene forekom hyppigst blant de fertile. Tre diagnoser ble diagnostisert hos mer enn 92% av kvinnene. Noen diagnoser forekom hyppigst hos infertile mens andre hos fertile.

Akupunkturpunktene Ki3, Liv3 og Sp6 ble valgt for nesten alle kvinnene, mens Cv4, Cv3 og St29 ble brukt hyppigst for infertile, og Sp3 hyppigst for fertile.

### **Konklusjon**

Resultatene viste utstrakt variasjon og svak inter-rater reliabilitet for TKM diagnoser.

Symptomene ble brukt inkonsekvent til å sette en diagnose. Liten enighet om diagnoser indikerer at akupunktørene følger sin egen diagnostiseringsprosess, selv om de har tilsvarende akupunkturutdanningsbakgrunn. Valg av akupunkturpunkter ser ut til å være nært relatert til valg av TKM diagnoser. Dataene indikerer at noen TKM diagnoser er tilstede og andre ikke i forbindelse med infertilitet. Disse resultatene er ikke konsistent med antagelsen om en forbindelse mellom åtte spesifikke TKM diagnoser og infertilitet, som hevdet i TKM lærebøker. Uoverensstemmelser og variabilitet innenfor TKM diagnoser kan svekke den individuelle tilpassede behandlingen, hvilket er en utfordring for klinisk praksis.

## Introduction

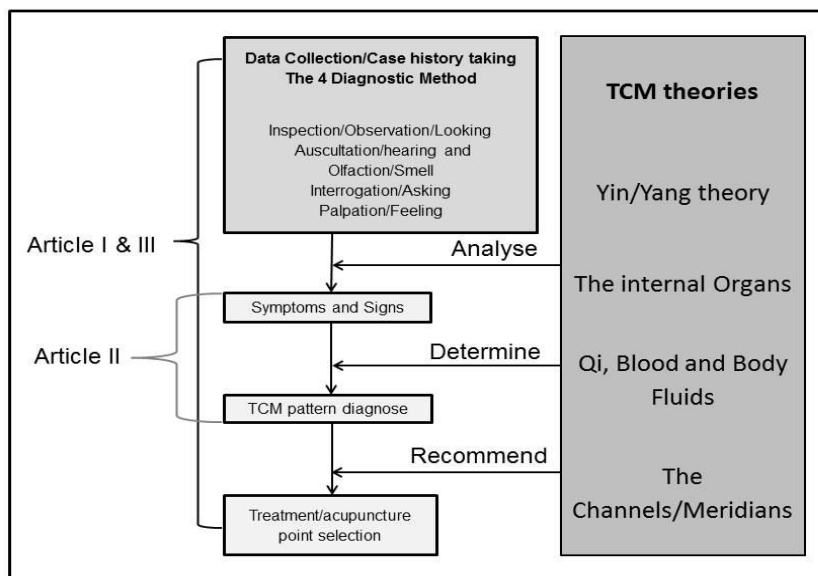
Acupuncture involves the insertion of needles into humans for remedial purposes, according to the World Health Organization's (WHO) definition.<sup>1</sup> Acupuncture has been a popular treatment method in Europe for many years.<sup>2</sup> About 10% of the Norwegian population reported in 2007 that they had used acupuncture over the past 12 months.<sup>3-5</sup> In Norway acupuncture is currently offered in private outpatient clinics and in the public health service,<sup>3;4</sup> and it is used as part of the management of a number of clinical conditions.<sup>6;6;7</sup> For instance, some hospitals recommend acupuncture in conjunction with in vitro fertilization (IVF).<sup>6;7</sup>

Several forms of acupuncture exist, and this study refers to acupuncture rooted in Traditional Chinese Medicine (TCM). This implies that the diagnostics and treatment regimen are based on TCM theories which constitute a framework quite different from the biomedical perspective.<sup>8;9</sup> As illustrated in Figure 1, the TCM theories explain how to collect data, interpret and understand the symptoms and signs in relation to one another, and use the information to set a TCM pattern. Finally, the TCM theories are used to determine the treatment according to the diagnosed TCM pattern, that is, which acupuncture points to use. Thus, individualised treatment according to the diagnosed TCM patterns is often regarded an important feature of TCM acupuncture.<sup>10-12</sup>

The TCM theories relevant to this study are briefly described in this section. The theories are presented in the current language used in contemporary TCM textbooks. Figure 1 shows where the theories are used in the diagnostic process.

Maciocia's work was chosen as the general theoretical basis for this thesis. Giovanni Maciocia is Italian and a world-renowned author and practitioner of Chinese Medicine (CM) who presented a structured overview of the TCM theories.<sup>8;9;15-18</sup> Maciocia studied TCM in mainland China. He introduced some of the first accounts of TCM theories into Western society, and made them accessible.<sup>19</sup> His books are used world-wide in the curricula in acupuncture education, including the Norwegian Acupuncture College (NAC). The acupuncturists who participated in this study were all familiar with Maciocia's presentation of TCM.

Figure 1. TCM theories used to analyse symptoms and signs to set a TCM pattern. The data are gathered through four examinations: Inspection, Hearing & Smelling, Asking and Feeling. With the aid of the TCM theories of pattern identification, the data are classified as interior or exterior, cold or heat, vacuity or repletion, and yin or yang to finally form the TCM pattern<sup>34</sup> and recommend the treatment.



## Development from old theories to modern TCM acupuncture

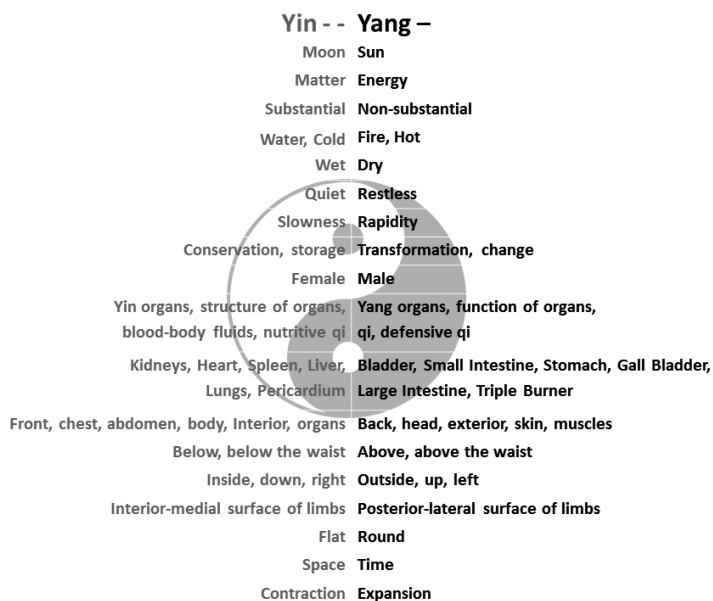
The theoretical foundation of TCM acupuncture can be traced back to the ancient Chinese texts, the *Huangdi neijing*, colloquially referred as *Neijing*.<sup>20</sup> These texts are dated between the late Warring States period (475–221 BC) and the early Han period (206 BC–220 AD).<sup>20;21</sup> The *Neijing* covers the diagnostic methods, and these basic theories are still guiding the TCM clinical practice.<sup>9;22</sup> Taylor and Scheid explains that later TCM textbooks, as for instance Maciocia, references to these old texts are based on excerpts of classical sources provided in the first national textbook, *the Outline of TCM* from 1958 and *Revised Outline of TCM* from 1972.<sup>19;20</sup>

### Yin and Yang

The Yin and Yang theory is the foundation of all the other TCM theories about the diagnosis and treatment of diseases.<sup>23</sup> The earliest reference to Yin and Yang is in the *Book of Changes*,

written around 1000 BC.<sup>9</sup> Originally, the concepts were used to indicate an object's relation to the sun, with the aspect facing the sun being Yang and warm, while that in the shadow being Yin and cold.<sup>24</sup> Figure 2 provides examples of the opposite features of Yin and Yang and illustrates the Yin and Yang symbol.

**Figure 2.** *Yin and Yang symbol, the Supreme Ultimate (Tai Ji), forming a circle to signify unity and complementarity. The white half represents the Yin part. The phenomena related to Yin are described on the left, and those related to Yang are described on the right. The black half represents Yang, and the white dot in it symbolises that Yang contains a seed of Yin and can change into the other and vice versa.<sup>9</sup> This figure was redrawn by Birkeflet, and inspired by Maciocia's work.<sup>9</sup>*



The theory was used in all fields to explain the theoretical issues; all natural phenomena were classified according to Yin and Yang during the tenth to eleventh centuries BC (the Shang and Zhou dynasties).<sup>9;24</sup> The Yin and Yang theory was introduced to the field of medicine by the *Neijing*, and it became the most important theory of CM.<sup>9;24</sup>

The foundation of the entire Yin and Yang energies of the body is attributed to the Kidneys. The Kidneys are significant to infertility since one of their functions is to govern reproduction, according to the TCM theories. Kidney Yang warms the body; hence, a decline in its functions will lead to symptoms such as aversion to cold, a whitish coating on the

tongue, listlessness and a weak pulse. These symptoms are used to set the diagnosis – Kidney Yang Deficiency – which is one of Maciocia’s eight infertility patterns.<sup>17</sup>

Yin and Yang are interdependent and in a constant stage of dynamic balance; they are not static but transform into each other. In an overheated body, an excess of Yang will produce sweat; then the water, Yin, will cool down the body to its normal temperature to recover its balance. The healing principle of TCM is thus to restore the body’s balance. Every phenomenon in the universe alternates through cyclical movements of peaks and bases.<sup>9</sup> Yang is in its purest and most refined form, totally immaterial and corresponds to pure energy. In its coarsest and densest form, Yin is totally material and corresponds to matter. Energy and matter are two states of a continuum, with an infinite possible number of states of aggregation.<sup>9</sup> Body regions, organs, symptoms and signs are classified according to the nature of Yin and Yang,<sup>1,9</sup> that is, the internal organs are classified according to Yin Organs and Yang Organs, as shown in Figure 2.

### **Internal organs – Zhang-Fu Organs**

The internal organs – the Zhang-Fu Organs – consist of the six Zhang or Yin Organs and the six Fu or Yang Organs. The Yin Organs are; Kidney, Heart, Spleen, Liver, Lung and Pericardium. The Yin Organs store the Essences extracted from food and use them to produce Vital Substances. The six Fu or Yang Organs are the Urinary Bladder, Small Intestine, Stomach, Gall Bladder, Large Intestine and Triple Burner. The Yang Organs receive and move food and drinks and then digest and excrete waste products.<sup>9</sup> The Triple Burner is defined as having its own function and representing the junction of organs situated in the upper, middle and lower burner/trunk and their functions.<sup>9,25-27</sup> The Triple Burner is not recognised in biomedicine; it does not seem to have an anatomical existence. This illustrates the different understanding guiding practices based on biomedicine and TCM, as well as the challenges in comparing and working across the two perspectives.

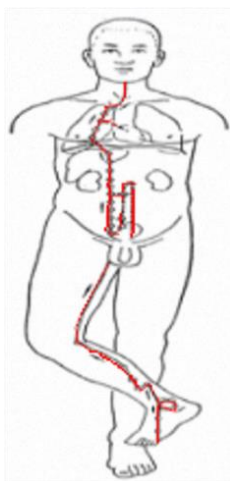
Chinese drawings of the organs, produced during or after the Song (960–1279) and Yuan dynasties (1271– 1368), deviate from the true structures of the human organs.<sup>25-27</sup> The human anatomy in *Neijing* was discredited as inconsistent and self-contradictory by Wang Qingren, who observed real human organs in 1831.<sup>28</sup> Despite the discredit and the influence of biomedicine, the old understanding about the organs is still guiding TCM clinical practice, which often refers to *Neijing*.

In *Neijing*, the knowledge of anatomy was based on metaphors. The structure and relations of organs were philosophical; they were concepts aimed at capturing the transformative processes animating the cosmos.<sup>25-27</sup> The body was compared to the perceived universe, and the metaphors related the organs to the functioning of the Chinese imperial government. The Zhang Organs played their dedicated roles as depots and the Fu Organs as palaces. The depots and palaces were linked by conduits, transportation channels or waterways that were essential for transportation and communication in China.<sup>26</sup> The transportation channels were often called meridians, and this term is used in the present thesis.

### **The meridians**

The meridians constitute an invisible duct system that combines with the blood vessels to comprise a network connecting the body parts to an organic whole.<sup>1;9;9;18;25</sup> Each organ is linked to a meridian, for example, the Kidney meridian shown in Figure 3. When the meridians are open, they can transport the Vital Substances between Zhang Organs/depots and Fu Organs/palaces. Qi, Essence, Body Fluids, Blood, Yin and Yang can flow freely through the meridians and nourish the organs properly; hence, they help make the human body healthy and fertile.<sup>9;25</sup>

**Figure 3.** *Illustration showing the course of the Kidney meridian, from the sole of the foot through the abdomen and deeper into the back, ending in the neck. The Kidney is a Yin Organ; consequently, the Kidney meridian is a Yin meridian. The figure was redrawn by Birkeflet (2014), inspired by Maciocia's work.<sup>9</sup>*





### **Vital Substances**

The Vital Substances produced and stored in the organs are Qi, Essence, Blood and other Body Fluids. Qi is a comprehensive concept; in this thesis, it is referred to as energy.<sup>9</sup> Qi refers to both the refined nutritive substance – the Nutritive Qi that nourishes the body and the mind – and its functional activities. The functional activities of the internal organs indicate the Liver Qi, and so on.<sup>9</sup> The Liver ensures the smooth flow of Qi in the whole body. The Stagnation of the Liver Qi can therefore obstruct the body's general circulation, including the circulation of the Nutritive Qi and the Vital Substances in the Uterus, consequently leading to infertility. Blocked nutrition to the Kidneys will thus affect their reproductive function.<sup>9</sup> Essence is a kind of energy with a fluid nature. It is stored in the Kidneys and affects reproduction, conception and pregnancy.<sup>9</sup> Sufficient and free-flowing Essence is thus important for fertility. Blood is regarded both as a form of Nutritive Qi and a part of Yin, as it nourishes and moistens the body.<sup>9</sup> Body Fluids indicate fluids of living organisms, the organic fluids, which pertain to Yin and have a nourishing function.<sup>9</sup> Consequently, the Vital Substances are essential for life and reproduction, as well as in their function to nourish the organs. The nourishment can be impaired and may accordingly show symptoms and signs from the affected organs.

### **Identifying imbalances in nourishment**

The theory of identifying patterns according to the Qi, Blood and Body Fluids and the theory of the internal organs are used to determine imbalances in nourishment. These theories are based on the interpretation of the symptoms and signs arising when the Qi, Blood, Yin and Yang of the internal organs are imbalanced. For instance, symptoms such as vaginal discharge, prolapse of the Uterus, weakness of the back, a pale tongue and deep weak pulse can indicate a weak Kidney Yang.<sup>9</sup> These symptoms of Kidney imbalances differ from the biomedical understanding of Kidney diseases, highlighting the contrast in the knowledge of the organs held by the two branches of medicine.

The energies in the meridians can be obstructed, be in excess or deficient and consequently create malnourishment or imbalances in the organs they serve.<sup>9</sup> The meridian theory can be used to identify imbalances in the body. Symptoms, pain or weakness in an area (e.g., the back) can be attributed to the organ associated with the meridian that runs through the affected area. The Kidney meridian goes from the sole of the foot towards the sacrum and

the lumbar spine and enters the Kidney. A symptom along the Kidney meridian area indicates a weak Kidney energy, which is associated with infertility.<sup>29</sup> Weakness in the Kidney energy could for instance be due to the climate and seasons, or a large number of other factors that could have impacts on the transportation channels and the depots and palaces.<sup>25</sup> Hence, a wide range of factors is believed to cause imbalances in the body.

### **Pathological factors**

According to the TCM theories, anger, irritability and frustration are pathological factors that affect the body, especially the Liver, and may cause Liver Qi Stagnation, which in turn may reduce uterine nourishment and cause infertility.<sup>9</sup> Fear depletes the Kidney Qi, which is essential for reproduction.<sup>9</sup> The external causes of diseases are the climatic factors, such as Wind, Cold and Damp, that can enter the body and cause internal imbalances. Internal Dampness can also be created by the intake of too much greasy, fried or fatty food, and the formation of Dampness can result in ovary obstructions. Too much cold foods (foods that give cold energy), including banana, strawberry, tomato, spinach and salad, can cause internal Cold. Cold and Damp that settle in the Uterus reduce circulation and nourishment and consequently prevent pregnancy.<sup>9</sup> Too much alcohol and spices may cause Heat to rise and subsequently dry up nourishment and disturb circulation. A generally poor health condition is related to weakness of the Kidney energy and the Vital Substances, whose strength and balanced functioning are required for pregnancy.<sup>9</sup>

Age is one factor that may impair the Kidney's energy for reproduction and thus may affect fertility. When a woman reaches the age of about 35, her Kidney Yang begins to weaken.<sup>30</sup> Smoking and drinking coffee also deplete the Kidney's energy and consequently reduce fertility, according to the TCM theories.<sup>31;32</sup> Body piercing, artificially induced abortions and surgery are considered traumas. Scarring and cutting of the meridians in the lower abdomen can result in Stasis of the Blood in the Uterus and consequently, infertility.<sup>30</sup> Oral birth control pills cause Stasis of the Blood, damage the Kidneys and obstruct the Spleen Qi, which lead to internal Phlegm and Dampness. Consequently, the energy and blood circulation to the Uterus will stagnate, and the Essence supplies will become insufficient, hence weakening fertility.<sup>30</sup> These examples show how current life style and modern human life can be linked to changes as described by TCM theory.

## **Pattern differentiation – fundamental principle of TCM diagnostics**

A pattern is the manifestation of a human illness; it is the diagnostic conclusion, drawn solely from clinical observations.<sup>1;9</sup> The diagnostic methodology, pattern differentiation or pattern identification (*Bianzheng Lunzhi*) is fundamental for TCM and regarded as a requirement for good clinical practice.<sup>20;29</sup> It means identifying and differentiating the imbalance (the disease), setting the pattern (the diagnosis) and determining the treatment. Hence, the diagnostic process does not differ from that of biomedicine, but the terms and theoretical foundation differ.

The pattern concept is different from the biomedical concept of diagnosis and disease;<sup>1</sup> as it is well recognized that a biomedically defined disease may be associated with several different patterns.<sup>33</sup> Distinctions among symptoms, patterns and diseases, as well as ideas about diseases, came to China by means of western medical knowledge and were never clearly defined before the 1950s.<sup>19;20</sup> A pattern indicates the nature, location or cause of an imbalance<sup>34</sup> and provides directions for treatment method and strategy.<sup>9</sup> In clinical acupuncture practice, the whole process of TCM pattern differentiation is guided by TCM theories (cf. Figure 1). The theories are applied to collect the data (symptoms and signs), interpret and analyse these in relation to one another, lead to diagnoses (the patterns) and finally, determine the treatment.<sup>9;24;34</sup> Thus, individualised treatment according to the diagnosed patterns is an important feature of TCM acupuncture.<sup>10-12</sup> To ensure consistent and optimal treatment, high reliability in identifying the TCM patterns is essential.<sup>14</sup>

The patterns are identified by inspection, auscultation & olfaction, interrogation and palpation (cf. Figure 1). Inspection refers to the examination of the patient's body posture, figure, vitality (brightness) and colour. The texture and quality of the patient's discharges are also considered, based on to the patient's description. Auscultation includes the examination of the sound of the patient's speech, respiration and cough. Olfaction is used to characterise the smell of the patient's body, secretion and excreta. Interrogation means interviewing the patient about other symptoms and signs. Palpation includes the examination of the radial pulse and palpation of tender areas if needed.<sup>17</sup> The data are based on subjective assessment, and biomarkers are not used. Hence, the data are vulnerable to individual interpretation, both when collected and analysed according to the TCM theories, which can be understood differently.<sup>17</sup>

**Treatment principles**

The basic idea of acupuncture treatment is to balance the body energies according to the identified imbalances. It aims to dissolve stagnated energies, reduce excessive patterns or strengthen deficiency patterns. This goal is achieved by identifying appropriate acupuncture points and insert needles into those areas. The needles are left in place for 20 minutes. The acupuncture points are located in the meridians. Longitudinally, each meridian covers several acupuncture points. Each acupuncture point has a specific function, for example, to regulate the Yin, Yang, Blood and/or Qi in different organ systems. When a needle is inserted into an acupuncture point, the Nutritive Qi is activated and supplies the actual area.<sup>10-12;35</sup> According to the TCM theories an effective treatment depends on the correctly diagnosed TCM pattern; hence, it is crucial to identify the correct pattern.

## Infertility and TCM

### TCM patterns related to infertility

According to the WHO, infertility is a disease.<sup>36</sup> Maciocia has proposed that the disease infertility can be categorized by eight different TCM patterns: Kidney Yang Deficiency, Kidney Yin Deficiency, Blood Deficiency, Cold in Uterus, Dampness in the Lower Burner, Blood Heat, Qi Stagnation and Stasis of Blood.<sup>17</sup>

After a thorough search, an additional 19 TCM textbooks about female infertility were identified.<sup>8;17;30;31;33;37-50</sup> These described altogether 130 different TCM patterns related to female infertility. The body of literature offers a comprehensive picture, and it seems that each author has more or less formed his or her own understanding of TCM patterns related to female infertility. The TCM textbooks indicate an association between certain TCM patterns and infertility. Beyond this, the books provided poor presentations of the empirical findings on which their interpretations were based. The presentations were mainly based on examples and anecdotes. Furthermore, the authors did not discuss the logical defects and theoretical weaknesses or the lack of consistency across the textbooks.<sup>8;17;30;31;33;37-50</sup> The low consensus and wide variations among authors reveal a lack of common understanding of the relationship between TCM patterns and infertility.<sup>8;17;30;31;33;37-50</sup> As the texts by Maciocia are widely used, the current study also applied his system for categorizing patterns.

The evolution from the understanding of childlessness in the old texts to the creation of eight TCM patterns related to infertility was not explicitly described by Maciocia. He states that the oldest Chinese text mentioning the treatment for childlessness was from the Warring States period (475–221 BC). This text describes medicinal plants for treating childlessness.<sup>17</sup> Moreover, by searching Maciocia's texts available in English, it has not been possible for me to find his implicit reasoning of how the TCM patterns became related to the biomedical diagnosis of infertility. To gain a deeper understanding of this relationship, I have therefore explored the recent literature about TCM and the influence of biomedical science.

## **Biomedical influence on the evolution of Chinese medicine**

Mao wanted to unite CM and biomedicine into a single, new and unique medicine.<sup>23;28</sup>

Chinese Medicine was criticised for not being scientific, and it had to be better integrated with modern science to raise CM to a level comparable with that of western medicine.<sup>23;28</sup>

Transforming CM was not meant to replace the medical tradition but to upgrade CM to the level of a scientific theory.<sup>51</sup> The unification process had major challenges.<sup>19</sup> The old understanding was not rejected, but an attempt was made to incorporate it with biomedicine.<sup>9</sup> The old Chinese medical concepts were adjusted to fit Marxist dialectics, western science and western biomedicine.<sup>25</sup>

In the process of systematising the CM theories in Communist China, the theories had to be politically correct. All scientific investigation was to be guided by the Communist philosophy.<sup>9;29</sup> To comply with the requirements of the prevailing political power, military metaphors for the organs were introduced. For instance, the Liver was associated with the attributes of a general<sup>19</sup> – one who possesses courage, derives strategies, plans people's lives and is resolute.<sup>9</sup> Militaristic notions were applied to the explanation of health and illness through the basic precepts of attacking and invading the body.<sup>19</sup> The body could be attacked by pathogenic factors and be exposed to imbalances in the Vital Substances, such as Yin and Yang imbalances.<sup>9</sup>

## **Pattern identification as a tool to integrate CM and western medicine**

The diagnostic system of pattern recognition, Bianzheng Lunzhi, has provided a mechanism to integrate CM and western medicine. Bianzheng Lunzhi had connections to ancient CM scholarship<sup>29</sup> and differentiated CM from western medicine, but it provided the theoretical tool and practical method for uniting the two different medical traditions.<sup>37</sup> Chinese Medicine borrowed theories and concepts from biomedicine to integrate the two medical systems. The concepts rarely corresponded exactly, but CM referred to the ancient medical text and incorporated biomedicine.<sup>27</sup>

With the diagnostic system of pattern recognition, Bianzheng Lunzhi, equivalence between the concepts of disease (Bing) and pattern (Zheng) was established.<sup>27</sup> Bing loosely resembles the biomedical concept of disease, and was further sub-categorised into Zheng or a pattern.

A pattern was defined by a unique cluster of symptoms and signs.<sup>27</sup> Zheng, which loosely means symptom, was transformed into the new diagnostic category of a pattern.<sup>27</sup>

The TCM pattern was made analogous to the biomedical concept of a disease – the origin of the pattern-centred approach to Bianzheng Lunzhi.<sup>29</sup> Bianzheng Lunzhi had not been established as a theoretical system prior to the 1950s.<sup>20</sup> It was developed from 1949 as a central methodology of TCM.<sup>52</sup> The systematisation of diagnostics was emphasised as the fundamental principle of TCM from the early 1950s.<sup>20</sup>

Another initiative to close the gap between CM and biomedicine and create a scientific fundament involved developing schools and curricula.<sup>20;29</sup> Through the development of the textbooks in the 1960s, complex CM theories were reduced to more easily accessible TCM theories and integrated into a national set of TCM textbooks.<sup>19;29</sup> Modern clinical TCM textbooks were organised according to major biomedical disease categories.<sup>29</sup> Hence, several TCM patterns specific to a given biomedical disease were described,<sup>20</sup> for example, the eight TCM patterns related to infertility.<sup>17</sup>

This organisation creates an impression of a direct connection between the biomedical disease (infertility) and the TCM patterns.<sup>17</sup> Taylor argued that the biomedical system of classification of diseases is not applicable to the incorporation of TCM-pattern diagnoses.<sup>19</sup> It is emphasised that the relationship between a TCM pattern and a biomedical disease is indirect and complicated, rarely corresponding exactly,<sup>9;29</sup> and direct connections is impossible because the two systems were based on completely different principles.<sup>17</sup> Nonetheless, many TCM patterns have been correlated with major biomedical disease categories in TCM textbooks.<sup>19</sup>

### **One pattern – many diseases, one disease – many patterns**

A single pattern (e.g., Stasis of Blood) is described as related to many diseases for instance; infertility, abdominal pain diseases, ovarian cysts, dysmenorrhoea, cerebral stroke, sciatica, headaches and irritable bowel syndrome.<sup>16;17</sup> A reason for this could be that the pattern is diagnosed on the basis of general symptoms and signs, which are common in many biomedical diseases, hence it is possible to diagnose the same pattern in different diseases.

A disease or condition such as infertility can be differentiated by several patterns, for instance the eight patterns described by Maciocia. This does not imply that the patterns are

subcategories, rather they describe different aspects of energy balance according to TCM theory. Infertility can be associated with a number of symptoms, and thus provides a foundation for several patterns. In addition, a symptom can be used to diagnose a number of different patterns, depending on the presence and absence of other symptoms. For instance, scanty menstrual bleeding and dizziness are symptoms that Maciocia relates to the patterns Kidney Yin Deficiency, Kidney Yang Deficiency and Blood Deficiency.<sup>17</sup> The last two patterns also share depression and a pale tongue in common.<sup>17</sup>

Consequently, the eight patterns described by Maciocia are not unique to female infertility.<sup>17;20</sup> Actually, fertile and infertile conditions can share the same TCM patterns. Therefore, it is difficult to understand how the association between TCM patterns and a biomedical disease or female infertility could be possible. It has been explained by ascribing a specific fertility function to an organ, such as the Kidneys that are connected to the reproductive function.<sup>17</sup> However, according to the holistic understanding of TCM, the whole body is interconnected. The meridians connect all parts of the body to an organic whole; thus, a problem in one part will create a problem in another part.<sup>9</sup> Therefore, all parts are important for fertility, and any imbalance could lead to infertility. If specific patterns are associated with infertility, then more of them could be expected to be found among the infertile, compared to the fertile ones. In this thesis, I have tried to examine the relationships between TCM patterns and biomedical diseases.

### **Significance of the patterns**

Since the TCM diagnostic criteria are not strict, any conclusion is based on individual interpretation. According to Wang et al., a TCM pattern is a product of speculation.<sup>53</sup> They maintain that the process involve imagery thinking and depends on the curriculum, the practitioners' learning, medical experience, academic background and other factors.<sup>53</sup> For instance, having a clinical background as a nurse or physiotherapist might have an impact on the interpretation of symptoms. However, the importance of these factors is poorly understood, and empirical data are lacking.

Furthermore, for other professions, factors such as age, education and practice are shown to influence how practitioners understand and evaluate findings. There is no reason to believe that the picture is different for TCM practitioners. A TCM theory might be



interpreted differently depending on the acupuncturists' additional educational background and work experience. Many acupuncturists are trained as physiotherapists or nurses and thus have a clinical education. Other acupuncturists hold a bachelor degree in medicine which is a theoretical education in Norway. Such differences can influence on the understanding of the TCM patterns and recommendations for treatment and consequently, the inter-rater reliability in diagnostics. It therefore becomes important for the understanding of diagnostics to examine whether demographic variables can have impacts on the TCM patterns.

## **Aims of the present study**

The overall aim of the included studies was to provide more knowledge about the diagnostic process in acupuncture.

As treatment is based on the individual TCM patterns according to TCM theory, the quality of the treatment is heavily dependent on pattern diagnosis. This thesis therefore has an emphasis on exploring the reliability of the pattern diagnosis and in selection of acupuncture point.

If the classification into the eight TCM patterns is typical for infertility,<sup>17</sup> it should be expected that infertile women have a high prevalence of these TCM patterns compared with fertile women. We hypothesised that infertile and fertile women should differ in the prevalence of the TCM patterns described as typical for infertility. Furthermore, we expected differences in the prevalence of recommended acupuncture points between infertile and fertile women.

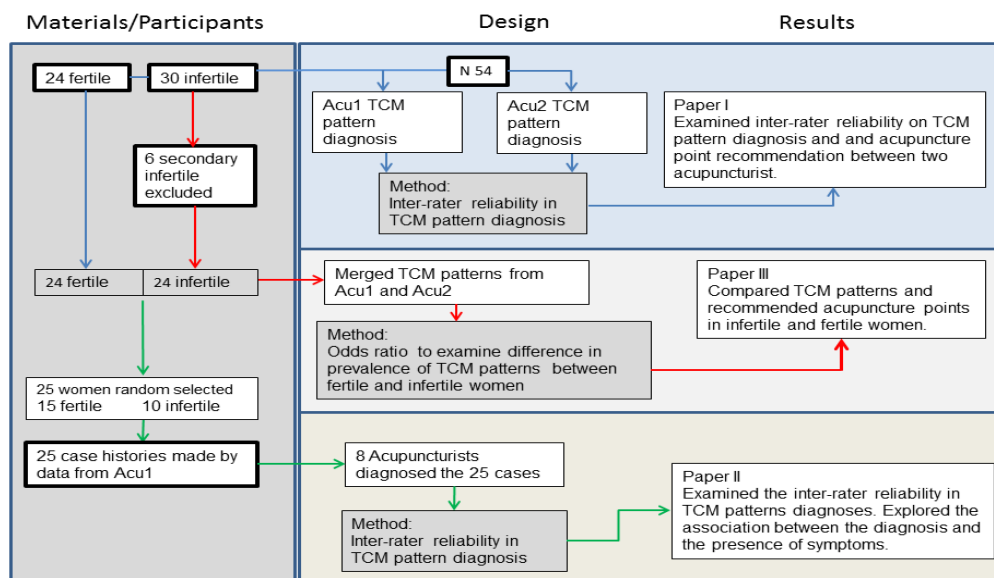
Separate studies (Papers I–III) were undertaken for the following reasons:

- I. Examine the inter-rater reliability of the TCM patterns and prescriptions for acupuncture points.
- II. Examine the inter-rater reliability of the TCM-pattern diagnoses made by eight acupuncturists for 25 case histories.
- III. Identify and compare the TCM patterns and recommended acupuncture points for infertile and fertile women.

## Study design, participants and methods

The inter-rater reliability of TCM patterns was examined with two methods. In Paper I, two acupuncturists, not blinded for fertility status, worked together in an ordinary TCM consultation. Based on the common data, collected during the consultation the two acupuncturists separately diagnosed each woman. In Paper II, the eight acupuncturists, who were blinded for fertility status, individually diagnosed (applying TCM) the 25 case histories made from the original sample in Paper I. Figure 4 illustrate the study design.

**Figure 4.** Schematic representation of the study design used in this thesis.



## Study Participants

The fertile and infertile women were recruited via advertisements on maternity care websites and posters displayed at physicians' offices. The women contacted the project leader by themselves and were included after a written consent. The participants from two groups, consisting of 30 infertile and 24 fertile women, were consecutively included and interviewed from September 2007 to October 2008.

The 24 fertile women were recruited by including those who had previously been spontaneously pregnant. They had delivered within the last 12 months prior to participating and were not pregnant at the time of their participation.

The infertile women were recruited from those included in an IVF programme who met the medical requirement for the infertility diagnosis – the failure to conceive following 12 months of unprotected intercourse.<sup>54</sup> Altogether 30 women with different and verified causes of infertility were recruited. This group consisted of 24 primary infertile (had never been pregnant) and 6 secondary infertile (had been pregnant previously) women. The six females with secondary infertility were excluded from Papers II and III (Figure 4) to ensure a largest possible difference between the fertile and infertile groups. The remaining 24 women consisted of 11 with unexplained infertility and 13 who reported endometriosis, polycystic ovarian syndrome and poor egg quality as causes of their infertility. All the 54 women were Norwegian-speaking, ranging from 24 to 42 years old. The average age for the infertile women was 33.4 years (24–41 years old); for the fertile women, it was 32.6 years (24–42 years old).

#### **Collecting data and taking case histories**

In clinical acupuncture practice, four diagnostic methods of inquiry are used to collect data: case history taking, palpation, inspection, observation and auscultation.<sup>17</sup> We chose research designs which minimized the variability in how the women's symptoms were presented to the acupuncturists. Two acupuncturists attended the interviews together and there were presented with exactly the same wording of how the women responded in the interview. They also could observe exactly the same body language. While one acupuncturist guided the interview, the other listened and made notes, with opportunities to ask additional questions. This ensured simultaneous access to identical information and minimised the potential for observational changes or bias to occur. Under the same light conditions, both acupuncturists examined each patient's tongue and felt her pulse simultaneously. The acupuncturists did not show or discuss their notes and findings with each other. To obtain data on the same items from all the participants, the interviewer followed a fixed questionnaire. This ensured that the information could be compared across all the women. For instance, if a woman denies having menstrual pain, she was still asked if she used painkillers for menstrual pain. This unmasked the fact that some women

experienced menstrual pain, and the answer was complete. This secured a comprehensive data collection that was important for diagnostics when young healthy women otherwise had few complaints. It also ensured a good basis for making case histories.

### Case histories

The data collected by Acu1 from interviews with 25 of the participants were used to make case histories. Each woman's case history contained all the information collected about her symptoms, signs, age and other observations. Information about being fertile or infertile was removed. A case history sample is shown in Figure 5.

**Figure 5.** Example of a case history. The case history reports all the symptoms and signs from one woman, with her pulse and tongue observed by one of the acupuncturists.

CASE 71; SYMPTOMS & SIGNS			
<b>General:</b> Female 27 years old. Smokes 15 cigarettes daily. Lifts heavy things. Bites nails, otherwise they crack because they are soft and thin. Has pain in the back and hips.			
<b>Appetite:</b> Alternates from good to bad, likes sweet and salty tastes. Eats irregularly and little, does not eat dinner every day. Feels thirsty and drinks about 1000 ml daily. <b>Stool:</b> Faeces one to two times daily, normal consistency/colour, one piece. Do have bothersome flatulence.			
<b>Sleep:</b> Well. <b>Temperature:</b> Normal temperate. Easily sweats under arms, big problem. <b>Gynaecology:</b> Menarche at 13 years. <b>Menstrual cycle</b> regular at every 28th days, lasts for 6 days, has normal flow. Menstruation has black colour, thick texture, it smells sour, no lumps. Abdominal and back pain 2-3 days before menstruation; a mild and dull pain is spreading like a belt around the waist. <b>Premenstrual tensions:</b> Headache, aching in the forehead.			
<b>Leucorrhoea:</b> Yellow, sticky, sour smell.			
PULSE		Right hand side	
Depth	LUNGE	SPLEEN	KIDNEY
Superficial	Wiry/slippy	Wiry/slippy	Not felt
Medio	Wiry/slippy	Wiry/slippy	Wiry/slippy
Deep	Wiry/slippy	Wiry/slippy	Wiry/slippy
PULSE		Left hand side	
Depth	HEART	LIVER	KIDNEY
Superficial	Slippy/Wiry	Slippy/Wiry	Not felt
Medio	Slippy/Wiry	Slippy/Wiry	Slippy/Wiry
Deep	Slippy/Wiry	Slippy/Wiry	Firm
TONGUE			
Tongue body is slightly swollen and has red colour, redder on the sides and in the centre. The tongue is dry. There are several, small transverse cracks in the middle of the tongue and on the tip. The coating is white, thin and dry.			

The 25 case histories were sent to eight acupuncturists for evaluation. The use of case histories (Paper II) provided an identical foundation for drawing conclusions in terms of TCM patterns and treatment.

## **Background of the acupuncturists**

All the acupuncturists had similar TCM education and were members of the Norwegian Acupuncture Association. Both acupuncturists being present at the consultation were educated at the Norwegian Acupuncture College, offering a bachelor's degree in TCM acupuncture. One of these two had 6 years of practical experience as acupuncturists. The other had 20 years' experience and also had the advanced course from Nanjing College of TCM in Nanjing, China.

Sixteen acupuncturists were invited to participate through phone and e-mail and 8 volunteered to participate in the study. They may be considered a strategic sample as they were selected to represent a broad group of acupuncturists with variations in geographic location, age and basic educational background (for instance, medical bachelor, physiotherapy or nursing). The demographic variables were examined as confounding variables in Paper II. The acupuncturist sample was small and not necessarily representative of the acupuncturist population. They held an equivalent of a bachelor's degree in TCM acupuncture, and their average age was 50, ranging from 33 to 59 years old. Three were males, and five were females. Two were physiotherapists, one was a registered nurse, and five had completed bachelor courses in medicine (MB). The acupuncturists who declined to participate were seven women and one man. Further information is not available for these acupuncturists.

## **Ethical considerations**

This study was approved by the Regional Ethical Committee; (Ref.nr.: S-09073d,2008/22929, 2009/2059). All the women received written and oral information about what participation in the study involved. A signed consent form was obtained from all. The woman behind the case history shown in Figure 5 gave additional written consent to use it in the paper and the thesis. The data from the women and the acupuncturists were stored together with an identification number, with the names and other personal identification removed. All analyses were made on anonymised data files. Demographic variables were omitted from table presentation since the acupuncturists could be identified.

In this study, the healthy volunteers were diagnosed by using TCM. None of the participants (neither the fertile nor the infertile women) was informed about their respective TCM

patterns; neither was anyone recommended for any treatment. The intention was to explore the TCM patterns existing in these women and how the patterns were related to their symptoms and signs, as well as to identify the relationship between the TCM patterns and biomedical conditions. We did not intend to determine whether some participants needed medical treatment for other complaints, which was also regarded as unnecessary since the women had not consulted the public health service for such a condition.

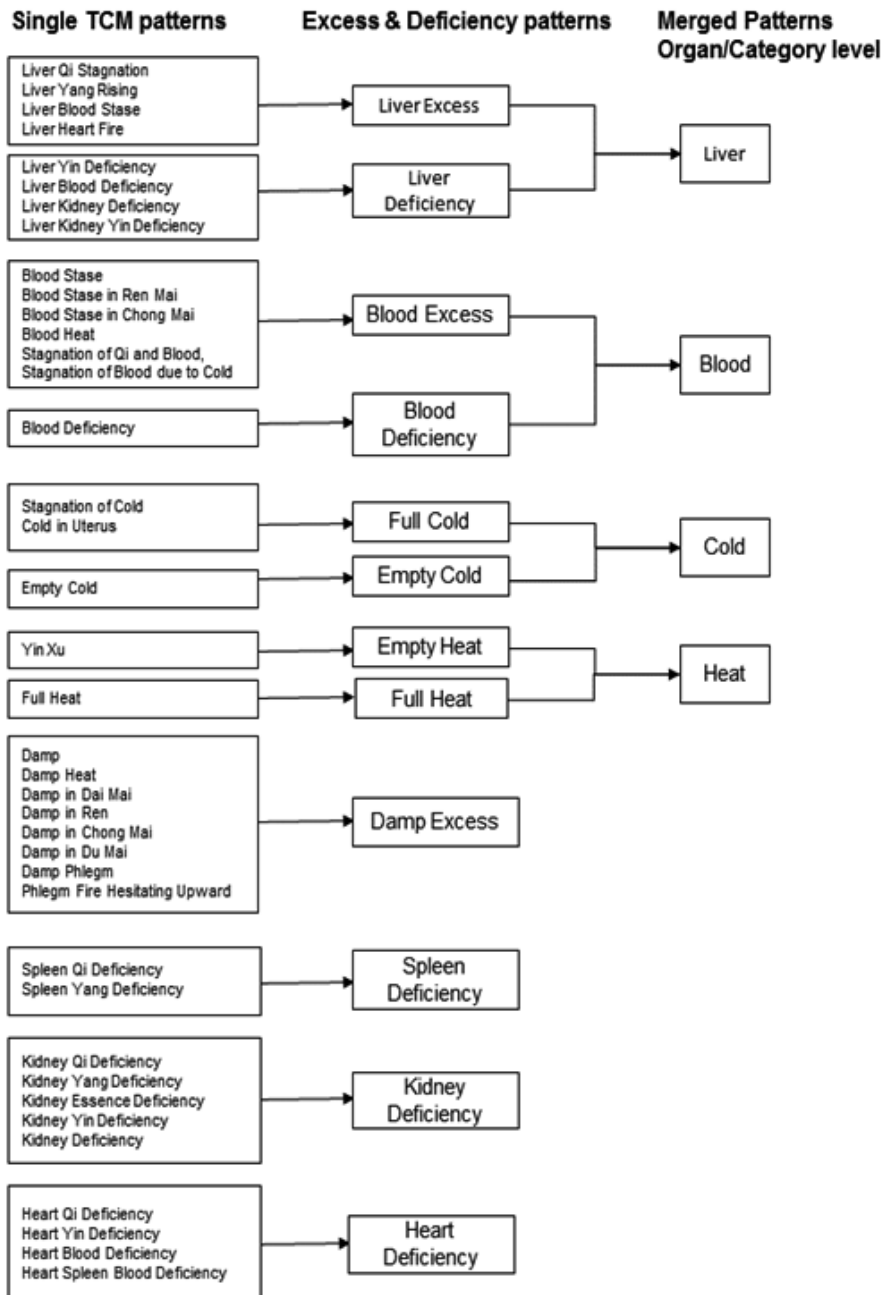
### **Data analysis**

As the acupuncturists used several TCM patterns, and some of them were used only once, the patterns were merged into groups. For each acupuncturist, single TCM patterns were merged according to their respective meridians' excess and deficiency patterns. Finally, the excess and deficiency patterns were merged for the organ level in Paper I (Figure 6) and Paper II (Figure 7).

The diagnosed TCM patterns were examined in Papers I and II. The frequencies of the patterns used in each participant's case and the average of the patterns that each acupuncturist used were analysed. Furthermore, the inter-rater agreements (kappa statistic) on the TCM patterns – the deficiency, excess and merged patterns – were examined for the two and eight acupuncturists in paper I and II, respectively. In Paper III, the differences in the prevalence of the TCM patterns between fertile and infertile women were examined, with the odds ratio as the effect measure. The frequencies of the TCM patterns distributed among the fertile and infertile participants were compared to Maciocia's eight TCM patterns related to female infertility.

The frequency of acupuncture point recommendations was examined in Paper I and III. The agreement on selecting acupuncture points and the uses of the meridians in treating the merged patterns were examined by inter-rater agreement (Paper II). The acupuncture point recommendations according to the fertility status were examined by binary logistic regression in Paper III. The symptoms and signs used to diagnose the TCM patterns were described in Paper II. For each case, the existing symptoms and signs were presented as frequencies and percentages. For each acupuncturist, the frequencies and percentages of the symptoms and signs used to set the TCM pattern of Liver Qi Stagnation were analysed.

Figure 6. Merging of single TCM patterns according to excess/deficiency for the organ level.





**Figure 7.** Merging of single TCM patterns via excess/deficiency patterns.

Single Patterns	Excess/Deficiency
Liver, Liver Qi Stagnation, Liver Qi Blood Stagnation, Liver Qi Blood Stase, Liver Qi Stagnation invading Stomach, Liver invading Lunge, Liver invading Stomach, Liver invading Spleen, Liver invading Heart, Liver Yang Rising, Liver Blood Stagnation, Liver Fire, Liver Excess	Liver Excess
Liver Blood Deficiency, Liver Heart Blood Deficiency, Liver Yin Deficiency, Liver Qi Deficiency	Liver Deficiency
Spleen Qi Deficiency, Spleen Blood Deficiency, Spleen Yin Deficiency, Spleen Yang Deficiency, Spleen Deficiency, Spleen and Dai Mai not holding	Spleen Deficiency
Stomach Heat, Stomach Qi Rebell	Stomach Excess
Stomach Qi Deficiency, Stomach Yin Deficiency, Stomach little moisture	Stomach Deficiency
Heart, Heart and Kidney not in Harmony, Heart Blood Stase, Heart Heat	Heart Excess
Heart Qi Deficiency, Heart Blood Deficiency, Heart Yin Deficiency, Heart Deficiency	Heart Deficiency
Heat, Heat Lower Jiao, Heat Middle Jiao, Heat Phlegm, Heat Phlegm Lower Jiao, Yang Burning, Fire Burning, Heat Intestine, Heat and Yang Rising	Heat Excess
Empty Heat, Yin Deficiency	Heat Deficiency
Blood Stase, Blood Stase Uterus, Blood Stase Cold, Blood Stase Cold Lower Jiao, Blood Stase Lower Jiao, Blood Stagnation, Blood Stagnation Uterus, Blood Stagnation Lower Jiao, Blood Heat	Blood Excess
Blood Deficiency	Blood Deficiency
Cold Stagnation Liver, Cold Uterus, Cold Invading Uterus, Cold Lower Jiao, Cold Stase	Cold Excess
Yang Deficiency	Empty Cold
Damp, Damp Lower Jiao, Damp Middle Jiao, Damp Heat, Damp Heat Uterus, Damp Heat Stomach, Dam Heat Liver and Gall Bladder, Damp Heat Liver Channel, Damp Heat Lower Jiao, Damp Heat Middle Jiao, Damp Cold, Damp Cold Lower Jiao, Damp Bi and Phlegm, Damp Retention, Damp Heat Retention, Damp Phlegm Retention, Damp Phlegm Heat Retention, Damp Bi, Damp Dai Mai, Damp Heat Chong Mai, Damp Heat Bladder, Damp Heat Intestine	Damp Excess
Phlegm, Phlegm Heat Yang Ming, Retention Phlegm, Retention Phlegm Damp Lower Jiao, Retention Phlegm Chong Mai, Phlegm Lower Jiao, Phlegm Kidney, Phlegm Dai Mai, Phlegm Lung, Phlegm Spleen, Phlegm Heart, Retention Phlegm Heat Lower Jiao, Retention Damp Phlegm Heat Lower Jiao, Retention Phlegm Sanjiao, Phlegm Heat	Phlegm Excess
Kidney Deficiency, Kidney Qi Deficiency, Kidney Yin Deficiency, Kidney Yang Deficiency, Kidney Yang Deficiency and Cold, Kidney Cold Disease, Kidney Jing Deficiency	Kidney Deficiency
Qi Stagnation	Qi Excess
Qi Deficiency	Qi Deficiency
Lung Qi Deficiency, Lung Yang Deficiency, Lung Deficiency, Lung Deficiency and Wei Bi Syndrome, Wei Deficiency	Lung Deficiency

## Statistical methods

Cohen's kappa statistic (k statistic) was used to assess the level of inter-rater agreement between the two acupuncturists (Paper I) and among the eight acupuncturists (Paper II). The k statistic measured the level of agreement beyond that expected by chance;<sup>55</sup> the k values were rated as follows: < 0.20 (poor agreement), 0.21–0.40 (fair), 0.41–0.60 (moderate), 0.61–0.80 (good) and 0.81–1.00 (very good).<sup>55</sup> The marginal totals for the 2 x 2 table were not balanced as the observed proportion of agreement is quite high, but the k value indicates a low level of reliability. This is a known paradox of the k statistic.<sup>56-58</sup> Since the k statistic alone is insufficient, we also report the maximum k value. The maximum agreement for k is 1.00, which is a perfect agreement, whereas 0 indicates no better than a chance agreement. Negative values show worse than a chance agreement.<sup>55</sup> The k statistic was calculated with SPSS 16.0 (Paper I) and SPSS 18.0 for Windows (Paper II). The 95% confidence interval (CI) and maximum kappa was calculated with DAG\_Stat (Paper I).<sup>59</sup>

In Paper II, the variations in the frequencies of the TCM patterns set by the eight acupuncturists were examined with respect to demographic covariates, and binary logistic regression was used. The variations in the use of the symptoms and signs to diagnose Liver Qi Stagnation were examined by descriptive analysis in Paper II.

In Paper III the TCM patterns set by the two acupuncturists were collapsed to produce one set of data. The same procedure was done for the acupuncture points. The merged data were used to examine the differences in the prevalence of the TCM patterns and acupuncture point recommendations for fertile and infertile women. The odds ratios, expressed as odds for fertility compared with infertility, were used as the effect measure. Additional analyses were performed to compare the results for each acupuncturist separately. To test the hypothesis regarding the absence of any difference in the odds ratio between fertile and infertile women, we used Fisher's exact test with mid-p correction as described by Lydersen et al.<sup>60</sup> To calculate the confidence intervals, we applied the Baptista–Pike mid-p method. These calculations were done in Stata and StatXact. Stratification by possible confounders was examined by the Mantel–Haenszel statistic. These calculations were performed with SPSS 18.0 for Windows.

## Main results

### Inter-rater reliability of TCM patterns (Papers I and II)

The two acupuncturists present in the consultation diagnosed a total of 39 TCM patterns (Paper I). Acu1 used 32 and Acu2 used 29 different patterns. On average, they used six and five patterns, respectively, for each participant. There was a large variation between the two acupuncturists regarding the TCM patterns they used for each woman.

Also when single TCM patterns were grouped and categorised as merged patterns ( Table 1), the results from the two acupuncturists showed poor to fair agreement based on k-values, whereas the max-k values indicate moderate to fair agreement. For the two most commonly used merged deficiency patterns – Spleen Deficiency (Max k = 0.34) and Kidney Deficiency (Max k = 0.25) – there was a fair agreement between the acupuncturists.

Even when the data were aggregated to organ level a striking variation was seen. For instance, Liver (Table 1) was diagnosed in 94% of the respondents, but still the two acupuncturists had a moderate agreement (Max k = 0.46).

Table 1. *Most used merged TCM patterns diagnosed by two acupuncturists, with mutual positive score, kappa (k), 95% CI and maximum kappa (Max k). n=54*

<b>Merged TCM Pattern</b>	<b>Positive score by Acu1 (n)</b>	<b>Positive score by Acu2 (n)</b>	<b>Mutual positive scores (n)</b>	<b>k</b>	<b>k (95% CI)</b>	<b>Max k</b>
<b>Liver Excess</b>	51	47	45	0.13	- 0.21 to 0.47	0.42
<b>Spleen Deficiency</b>	46	42	37	0.15	- 0.15 to 0.44	0.34
<b>Damp Excess</b>	21	27	14	0.22	- 0.03 to 0.48	0.25
<b>Kidney Deficiency</b>	44	37	30	-0.01	- 0.26 to 0.23	0.25
<b>Liver Deficiency</b>	11	20	5	0.08	- 0.17 to 0.33	0.25
<b>Blood Excess</b>	41	9	8	0.06	- 0.04 to 0.17	0.09
<b>Liver</b>	51	51	49	0.30	- 0.21 to 0.80	0.46

In total, 114 different single TCM patterns were used by the eight acupuncturists. Acu4 used the fewest patterns (63), and Acu7 used the most (203) patterns. They diagnosed on average 4 and 12 patterns on each case. Some of the patterns were only used once.

The TCM patterns set by the eight acupuncturists based on the 25 case histories varied greatly as shown in Table 2. The single patterns were grouped into the merged patterns, and

there were still wide variations among the eight acupuncturists (Table 2). Some acupuncturists did not diagnose Blood Excess or Heart Deficiency at all, whereas others used these on more than 10 of the 25 case histories (Table 2).

Table 2. Number of women with seven merged TCM patterns and the inter-rater reliability test on the seven merged TCM patterns diagnosed more than five times on average by the eight acupuncturists with kappa and 95% CI (Paper II). n=25

Merged TCM Patterns	Acu1 (n)	Acu2 (n)	Acu3 (n)	Acu4 (n)	Acu5 (n)	Acu6 (n)	Acu7 (n)	Acu8 (n)	k	95% CI
Liver Excess	23	25	17	19	25	21	24	23	0.014	0.000 to 0.036
Damp Excess	16	22	9	15	22	11	24	18	0.042	0.011 to 0.071
Blood Excess	7	1	0	1	3	17	17	7	0.144	0.065 to 0.222
Spleen Deficiency	21	23	12	8	21	17	25	23	0.029	0.001 to 0.056
Kidney Deficiency	15	15	7	6	14	19	22	15	0.043	0.001 to 0.076
Liver Deficiency	4	15	7	6	6	12	9	3	0.118	0.056 to 0.179
Heart Deficiency	7	11	4	0	7	1	19	6	0.179	0.095 to 0.263

Even for the most frequently used merged TCM patterns there was poor agreement among the eight acupuncturists ( $k < 0.20$ ), (Table 2).

#### **Impacts of demographic variables on the frequency of TCM patterns (Paper II)**

The age of the acupuncturists had no significant impact on the frequency of their diagnoses, but their gender, duration of practice and education (MB vs physiotherapist/nurse) were significantly associated with their use of merged TCM patterns ( $p < 0.001$ ). Moreover, the number of their work hours had a significant effect ( $p = 0.029$ ). Female acupuncturists and those with lengthy practical experience or longer work hours diagnosed fewer TCM patterns. Having a background as a nurse or a physiotherapist was associated with setting a higher number of TCM patterns compared to those who completed basic medical courses only.

#### **Variations in use of symptoms and signs to diagnose TCM patterns (Paper II)**

The use of symptoms and signs in diagnosing TCM patterns was examined in Paper II. A much used TCM pattern, Liver Qi Stagnation, was selected as a case to study because this

pattern was used on all 25 case histories by at least one acupuncturist. Altogether, 179 different symptoms and signs were used in the case histories, and 147 of these were reported by the acupuncturists to be used to diagnose Liver Qi Stagnation. Some acupuncturists used few symptoms, whereas others used several symptoms to make their diagnoses. Of the eight acupuncturists, Acu7 used the highest number of symptoms (3–37 per case), and Acu4 used the least (1–10 symptoms per case). On average these two used 15 and 3 symptoms per case, respectively.

There were wide variations among the acupuncturists in how symptoms were used in making their diagnoses. These variations were not related to how frequently the symptoms occurred. For instance, a wiry pulse (a pulse felt taut and string like a guitar string to the touch)<sup>33;61</sup> occurred in all the cases but was only used to diagnose some of them. One acupuncturist used a wiry pulse to diagnose 32% of the cases, whereas another acupuncturist used the condition to diagnose 84% of the cases. A less frequent symptom, red edges around the tongue, was never used in one acupuncturist's diagnoses, whereas another acupuncturist used it to diagnose all the 12 cases presenting with the symptom. Irregular menstruation, a rare symptom occurring in 4 cases, was not used by two acupuncturists, whereas one used it in 50% of the cases. The acupuncturists thus showed considerable intra- and inter-rater variations in their interpretations of the diagnostic meaning of symptoms and signs.

### **Treatment recommendations (Paper I)**

The recommended treatments included combinations of several acupuncture points for each woman. A total of 36 different acupuncture points were recommended by the two acupuncturists present in the consultation. Acu1 and Acu2 recommended 34 and 22 points, respectively. They recommended 8 and 6 points on average per woman.

The inter-rater reliability was examined concerning the recommendations on the single acupuncture points and merged points (meridians), according to the TCM patterns.

The two acupuncturists recommended Point LR3 for 51 and 47 women, respectively. In 44 of the women both acupuncturists recommended LR3. Moreover, points SP6 and KI3 were recommended for the majority of the women (Table 3). For these three acupuncture points, the maximum  $k$  indicated a fair to poor agreement. For the other points, the acupuncturists

differed to a large extent in the frequency of use, with a poor to fair agreement based on the maximum k.

**Table 3.** *Inter-rater reliability of the points most frequently recommended for the 54 women.*

Acupuncture point	Positive score Acu1 (n)	Positive scores Acu2 (n)	Mutual positive score	k (95% CI)	Max k
LR3	51	47	44	-0.16 to -0.01	0.39
SP6	35	45	29	-0.25 to 0.22	0.23
KI3	38	38	25	-0.40 to 0.09	0.18

When the acupuncture points were merged according to the meridians on which they were located, there was a moderate to fair agreement (maximum k) for the four most frequently used meridians – Liver, Kidney, Stomach and Spleen. Generally, the agreement was about the same as for the single points. The agreement was poor to fair for the Large Intestine, Conception Vessel, Heart, Gallbladder and Lung meridians.

The two acupuncturists' use of the meridians (merged acupuncture points) to treat a given merged TCM pattern was examined for agreement in Paper I. Since the Liver meridian was recommended for almost all women, it was found in combination with nearly all other patterns. A high agreement was observed for the Liver meridian in combination with all merged patterns. k and max k were not calculated for this pattern because either the agreement was 100 % or one of the acupuncturists had used the meridian in all the cases diagnosed with the pattern. For the other meridians, the two acupuncturists had a fair agreement on using them on the merged patterns.

### **TCM patterns in fertile and infertile women (Paper III)**

Altogether, 39 different TCM patterns were set after collapsing the data from the two acupuncturists present at the consultation. Some of these were only used once or a few times. Only the 15 patterns that were diagnosed in at least 5 of the 48 women were included for further analyses.

Five TCM patterns occurred significantly more frequently among the infertile women ( $p \leq 0.03$ ), whereas four TCM patterns were observed more frequently among the fertile women ( $p \leq 0.02$ ). Six patterns did not differ in frequency between the groups (Table 4).

Three of the five patterns with a higher frequency among the infertile participants; Kidney Yang Deficiency, Qi and Blood Stagnation, and Cold were among the eight patterns that Maciocia assumed typical for infertile women. On the other hand, two of Maciocia's eight patterns were more frequently used for the fertile women – Kidney Yin Deficiency and Damp.

**Table 4.** *The frequency of merged TCM patterns in the infertile and the fertile women. The five highlighted TCM patterns correspond to five of the eight patterns related to female infertility according to Maciocia.<sup>17</sup>*

TCM patterns	Infertile (n = 24), N(%)	Fertile (n = 24), N(%)	OR	CI	p
<b>Five TCM patterns that occurred more frequently among infertile participants</b>					
Kidney Yang Deficiency	15 (63)	5(21)	0.16	0.05 to 0.60	0.003
Qi and Blood Stagnation	9 (38)	2 (8)	0.15	0.03 to 0.79	0.03
Cold	11 (46)	3 (13)	0.17	0.05 to 0.67	0.01
Liver Blood Stasis	10 (42)	0	0	0.00 to 0.26	< 0.001
Spleen Yang Deficiency	10 (42)	3 (13)	0.20	0.05 to 0.81	0.02
<b>Four TCM patterns that occurred more frequently among fertile participants</b>					
Kidney Yin Deficiency	10 (42)	20 (83)	7.00	1.70 to 22.72	0.004
Damp	10 (42)	19 (79)	5.32	1.41 to 18.15	0.01
Liver Yang Rising	4 (17)	12 (50)	5.00	1.36 to 16.16	0.02
Heat	1 (4)	9 (38)	13.80	2.10 to 15.6	0.01
<b>TCM patterns that did not differ between both groups</b>					
Liver Qi Stagnation	24 (100)	23 (96)	0	0.00 to 9.00	0.50
Liver Blood Deficiency	13 (54)	9 (38)	0.51	0.17 to 1.65	0.27
Spleen Qi Deficiency	22 (92)	23 (96)	2.09	0.13 to 31.45	0.62
Kidney Deficiency	24 (100)	22 (92)	0	0.00 to 2.13	0.25
Blood Stasis	15 (63)	14 (58)	0.84	0.29 to 2.71	0.80
Stagnation of Blood due to Cold	6 (25)	2 (8)	0.27	0.05 to 1.35	0.15

Besides the TCM patterns that Maciocia related to infertility, Liver Blood Stasis and Spleen Yang Deficiency also occurred significantly more frequently among the infertile women. Among the fertile women, there were significantly more frequencies of Liver Yang Rising and Heat (Table 4). Three TCM patterns were common in both groups and occurred in at least 92% of both infertile and fertile women – Liver Qi Stagnation, Spleen Qi Deficiency and Kidney Deficiency.

### **Prevalence of acupuncture points for fertile and infertile (Paper III)**

The prevalence of the recommended acupuncture points for fertile and infertile women was compared. The acupuncture point recommendations by the two acupuncturists were

merged into one set of points for each woman. The distribution of recommended acupuncture points between infertile and fertile women was examined (Table 5). Three acupuncture points (CV4, CV3 and ST29) were recommended significantly more frequently for the infertile compared with the fertile women (0–0.07 odds ratio [OR],  $p \leq 0.03$ ). Point SP3 was chosen more frequently for fertile women (8.3 OR, 2.001–26.87 CI,  $p < 0.001$ ). Three points (KI3, LR3 and SP6) were recommended for almost all women in both groups. For the other points, no statistically significant differences were found between infertile and fertile women.

**Table 5.** *The frequency of acupuncture points recommended for infertile and fertile women. Frequency and percentage given for the two groups, with OR values, 95% CIs and p values. N=24 for both groups*

Acupuncture points	Infertile N (%)	Fertile N (%)	OR	CI	P
CV4	20 (83)	6 (25)	0.07	0.02 to 0.30	< 0.001
CV3	5 (21)	0	0	0.00 to 0.64	0.03
SP3	9 (38)	20 (83)	8.33	2.01 to 26.87	< 0.001
ST29	16 (67)	2 (8)	0.05	0.01 to 0.23	< 0.001



## Discussion

We found fair to moderate inter-rater reliability on the TCM patterns diagnosed by the two acupuncturists who joined the consultation together. On the other hand, there was poor agreement among the eight acupuncturists who diagnosed standardised information in the form of case histories. Altogether, the participating acupuncturists diagnosed 114 different patterns, with a wide variation between acupuncturists with respect to the number of patterns. Demographic variables among the acupuncturists (gender, work experience and working hours, and additional education) had an impact on the number of diagnosed TCM patterns.

The 179 different symptoms and signs that described the case histories were used differently and inconsequently when diagnosing the TCM patterns. Of the eight TCM patterns that Maciocia assumed typical for female infertility, three patterns occurred significantly more frequently among infertile women, whereas two of the patterns had a higher prevalence among the fertile women. Inter-rater reliability of the recommended treatments was fair to poor.

### Low inter-rater agreement on TCM patterns

The low inter-rater reliability found in the present thesis confirmed the conclusions in a review of reliability studies of TCM patterns.<sup>63</sup> The individual studies of reliability was carried out with participants experiencing a wide range of health problems, including temporomandibular joint disorder,<sup>64</sup> hypercholesterolemia,<sup>65</sup> rheumatoid arthritis,<sup>66;67</sup> low back pain,<sup>68;69</sup> stroke patients,<sup>70</sup> osteoarthritis of the knee,<sup>71</sup> ketamine abuse,<sup>72</sup> prediabetes,<sup>73</sup> tongue diagnosis,<sup>74</sup> pulse characteristics<sup>75</sup> and irritable bowel syndrome.<sup>76</sup> Six of the 13 studies did not report any statistical analysis as a foundation for their conclusions. The remaining 8 studies used kappa-statistics or Kendall's correlation and they all showed poor to fair agreement. Hence, it seems that low inter-rater reliability can be a general problem with TCM patterns irrespective of the health problem involved.

A common understanding among acupuncturists is that TCM cannot be analysed in the same way as western medicine including application of biostatistical methods. Based on a review of papers written in both Chinese and English, Shea found that a minority of the authors argued against the use of Western science and methods to analyse TCM

phenomena.<sup>77</sup> The basis for this view was that the assumptions entailed in the Western approaches are incompatible with TCM theories. The differences of importance include a wide variety of topics, including personal experience, the therapeutic encounter, the subjective basis of diagnostic procedures, holism and complexity. However, none of these apply as limitations in exploring reliability of diagnostic conclusions. There might be limitations related to the adequacy of the statistical methods for a given data set. This is however a general problem, and not a specific problem for TCM. In the present study we have used appropriate statistical methods for our data. The analyses are independent of theory, and do not provide information about the TCM theory. We therefore maintain that the conclusion of poor reliability is based on a valid analytical process.

### **Factors that could cause variations in diagnosis**

The variations in the diagnostics could be owed to weaknesses in the TCM theories and the factors in the TCM diagnostic process, both in the collection and interpretation of the data. In the data collection phase, subjective data are gathered. Determining the appearance of the tongue and describing how the pulse is felt could be perceived differently by different acupuncturists and result in varying TCM patterns.

The individual variations among the acupuncturists with regard to how they perceive a symptom, is thus one possible cause of variation even when they are presented with the symptoms simultaneously as in Paper I. Another cause of variation is related to the design of studies. In several studies, the participants have been examined by the acupuncturists in sequence, and hence the symptoms and signs may vary between the consultations.<sup>65;71;73;78;79</sup> In these situations, variation between acupuncturists does not necessarily reflect inter-rater reliability, but rather expresses the variability in the symptoms and signs. For instance, in two studies the tongue picture and pulse quality was assessed in separate consultations and the result was a large variation<sup>14;63</sup> Interpretation of this finding as low inter-rater reliability must be made with caution. In the present study we either used data from simultaneous observation (Paper I) or from written case histories (Paper II). Hence, we can exclude any shifts in symptoms and signs as a cause for the variability reported.

Su underlined the importance of collecting adequate clinical information. The collected data should undergo quality assurance through a TCM questionnaire.<sup>80</sup> In our study, we used a questionnaire to ensure that all the participants were asked identical questions reflecting clinical practice. Although we collected adequate and identical clinical information, the common finding was a low agreement on the TCM patterns. This indicated that the acupuncturists could also have interpreted the given set of data differently. This was confirmed by the eight acupuncturists who interpreted descriptive data in the form of written case histories, which resulted in a low inter-rater agreement on the diagnoses. The eight acupuncturists (Paper II) were blinded for the biomedical diagnosis. Thus, they were protected from confirmation bias; they had to search for evidence to confirm a diagnosis rather than refute it.<sup>81</sup> Since the eight acupuncturists interpreted exactly the same data, they must have understood and interpreted the data differently, concluding with different TCM patterns. Zhang et al. emphasize that the entire TCM diagnostic process relies on the practitioners' subjective observations and interpretation of the data. This can lead to inconsistent diagnoses, hence bringing the validity of TCM into question as a medical system.<sup>82</sup>

All symptoms and signs are considered to provide an overall picture of the whole person.<sup>10</sup> Comprehensive information, including the patient's general characteristics, body characteristics and shape, behaviours, personality, emotional condition, psychological state and interaction with the environmental climate,<sup>83</sup> is gathered and woven together until it forms a specific pattern of disharmony.<sup>11</sup> Even body piercing, surgical scars, birth control pills and diet can be used to make a diagnosis.

Skjelstad confirmed that if only a few of many symptoms would be needed to make a diagnosis, there would be a lot of combination possibilities. There would be a large variation in which symptoms to use for the diagnosis. Two patients with the same diagnosis could have a few symptoms in common.<sup>84</sup> This could pose a significant problem for our results, where a lot of symptoms were used for the diagnoses. In total, 147 (82%) of the 179 different symptoms that described the cases, both rare and frequently occurring, were used and interpreted as significant in diagnosing Liver Qi Stagnation. With the lack of a standardised diagnostic procedure, the acupuncturists could decide by themselves which symptoms to use. Hence, a large number of symptoms were used to diagnose Liver Qi

Stagnation, as discussed in Paper II. For instance, the symptom of red edges around the tongue was never used by one acupuncturist but was used in all cases by another. This underlines the fact that each acupuncturist's personal interpretation of the symptoms and signs influenced his or her decision. The symptoms were also used inconsistently by the individual acupuncturists, who diagnosed just some of the cases with the symptoms. For instance, a wiry pulse that occurred in all cases was used to diagnose just some of them, ranging from 32% to 84%. Consequently, irregular handling of data causes an uncertain diagnostic situation and increases the possibility of a low inter-rater agreement on the diagnoses. Hence, agreement on how to conclude with regard to TCM patterns requires clear guidelines and standards for handling the symptoms and signs.

The interpretation of data could also be affected by conflicting symptoms. For instance, one acupuncturist responded that too many cases were presenting with a wiry pulse; in fact, all the cases were described as such. Different from a clinical practice situation, the sample comprised a homogeneous group of relatively young women. Hence, a common symptom such as a wiry pulse would appear evident. A wiry pulse denotes a forceful type, an indication of a Liver Excess condition.<sup>9;33</sup> If coexisting symptoms indicated a deficiency pattern, there would be a contradiction. Maciocia confirmed the presence of contradictions in the TCM theories.<sup>17</sup> Some of the acupuncturists could consequently have placed less confidence in the wiry pulse and perceived it as an observation error since they did not feel the pulse themselves. Hence, it could result in different interpretations and conclusions. The acupuncturists were not asked to explain their evaluation of the data and how they handled conflicting/contradictory symptoms and signs. Table 6 illustrates the examination of a combination of many symptoms, providing an example of a situation that contributes to different conclusions. For instance, according to Maciocia, a pale tongue is a symptom of a Cold pattern. If the tongue is also dry, it could be a sign of Blood Deficiency, whereas if it is pale and wet, it could indicate Yang Deficiency.<sup>9</sup>

**Table 6.** *Maciocia's tongue diagnosis. It gives an example of how one symptom/sign can be interpreted differently, depending on coexisting symptoms. If the tongue is pale, it can indicate a Cold pattern, while a pale and dry tongue can be due to Blood Deficiency. A tongue that is both pale and wet points to a Yang Deficiency pattern.*<sup>9,15</sup>

Symptoms and signs	Co-symptoms	TCM patterns/diagnosis
Pale Tongue	none	Cold
Pale Tongue	and Dry Tongue	Blood Deficiency
Pale Tongue	and Wet Tongue	Yang Deficiency

In clinical practice, the acupuncturists can ask the patients to explain conflicting information to create a clearer picture. Compared to clinical practice, the analysis of a case history is an artificial situation. Our study provided no opportunity for the acupuncturists assessing the case histories to examine and touch the patients or observe their body language, complexion, voice, behaviour and so on. It was also not possible to supplement the available information by asking extra questions. These sources of information were reported as missing by some acupuncturists (Paper II), reflecting the idea that they could have drawn different conclusions if they had access to supplementary information. On the other hand, when the two acupuncturists (Paper I) were both present during the interviews and had access to all the information, low inter-rater reliability on the TCM patterns was also found. Hence, the completeness of the information, including the data on fertile/infertile conditions, did not seem to improve the agreement on the diagnosis.

As emphasised by Zhang, the background of the practitioners can influence their decisions regarding diagnosis and treatment.<sup>82</sup> As in the present thesis, Zhang et al. also found that a given set of symptoms was interpreted differently leading to a low agreement on TCM diagnosis. The present thesis showed that gender, experience, work hours and education had an impact on diagnosis and treatment choice. Female acupuncturists and those with lengthy practical experience and longer work hours diagnosed fewer TCM patterns. Acupuncturists with a background in nursing or physiotherapy diagnosed more TCM patterns than MB-educated acupuncturists. The MB acupuncturists had a theoretical education as a background before their acupuncture training, whereas the physiotherapists and nurses had clinical education and experience. Larger studies are needed to clarify if and

how the various educational background results in different practice in diagnosis setting and treatment choice.

Andersson et al.<sup>85</sup> confirmed Benner's theory<sup>86</sup> and concluded that nurses with extensive experience and specialised education reasoned differently than novice nurses. Experience, combined with further education, was important for developing professional competence. Becoming an expert required authentic situations that could not be offered in school. Expert nurses handled the data based on their own underlying previous knowledge and personal experience.<sup>85;86</sup> The acupuncturists could have applied their earlier personal experiences to acupuncture practice, which could influence how they considered the symptoms and signs.

### **Impacts of TCM theories on diagnosis**

The symptoms and signs are poorly specified or defined in the TCM theories. Hence, there is lack of a systematic and standardised way to interpret the observations. Diffuse definitions can cause disagreements. For instance, the symptom of 'heavy periods' is defined as more than the normal amount of menstrual bleeding.<sup>17</sup> Determining what is normal is a subjective task that could give different answers and consequently, various TCM patterns.

There is also some confusion in distinguishing between symptoms, TCM patterns and disease. For instance, heavy period is stated as a symptom of several patterns i.e. Kidney Yin Deficiency. Heavy period is also classified as a biomedical disease; a menstrual irregularity that is subdivided into three patterns: Qi Deficiency, Blood Heat and Blood Stasis.<sup>17</sup> Classifying the same term under different categories and giving vague definitions of patterns and symptoms can lead to a variable and individual understanding and different conclusions.

Maciocia underlines that the symptoms and signs described in TCM books are those expressed by Chinese patients. Western patients express the same symptoms in a different way.<sup>9</sup> Ramp et al. found significant differences in how specific symptoms were presented by German and Chinese menopausal women.<sup>87</sup> According to Scheid et al., the TCM literature explains menopause symptoms according to a biomedical understanding and ignores viewing the symptoms in relation to other factors, such as culture, diet and place of residence.<sup>88</sup> Local variations in specific symptoms are not emphasised in TCM textbooks.<sup>17</sup> These variations could therefore have contributed to misdiagnosis, in that symptoms could have been interpreted in an attempt to fit the pattern described in a textbook.

## TCM patterns in fertile and infertile women

Maciocia organise the infertility patterns according to biomedical diseases in his TCM textbook.<sup>17</sup> The eight patterns are related to female infertility by their attributes, explained through TCM theory and not by empiric results. For instance the Kidneys are attributed to reproduction, the Liver circulates blood and energy through the Uterus. Imbalances in the functions of these Organs can lead to infertility.<sup>17</sup> The imbalances are diagnosed from numerous and common symptoms and signs and there are general characteristics of patterns.<sup>9;17</sup> When the same symptoms and signs are used to make different diagnoses, Skjelstad explains that several disorders can appear, which could increase the number of a person's diagnoses.<sup>84</sup> A large number of diagnoses were made in our study. To diagnose many patterns are not regarded as problematic in TCM. A well-known saying is that one disease can have many patterns, and the same TCM pattern can exist in different diseases.<sup>9;19;20</sup> We confirmed that one disease, infertility, had a lot of patterns. The same was found in a study of menopausal women.<sup>89</sup> Furthermore, the results also show that some TCM patterns occurred quite frequently in both groups. The patterns were not unique to infertility; although three patterns associated with infertility were found in the infertile women, two were found among the fertile women.

The same patterns can exist in different diseases; the patterns diagnosed in paper III were also diagnosed in Borud et al.'s study on menopause.<sup>17</sup> The common pattern in infertile, fertile and menopause women<sup>13</sup> were Kidney Yin Deficiency, Kidney Yang Deficiency, Liver Qi Stagnation, Liver Yang Rising and Stasis of Blood. These patterns can be diagnosed on basis of general complaints. Borud et al. measured change in mid-aged women's emotional and physical health. Some of the symptoms measured were depressed mood, somatic and vasomotor symptoms, memory/concentration, anxiety/fears and sleep problems.<sup>89</sup> Equivalent symptoms were seen in the fertile and infertile, paper I. The symptoms are not unique to menopause/infertility, and may exist in other conditions. Hence, the same patterns can be diagnosed in different diseases, but it will be a fallacy to conclude that the patterns induce menopause or infertility.

General characteristics and common symptoms of both fertile and infertile women can give the same TCM pattern. Consequently, the diagnostic specificity would be weakened, and the diagnoses could lose discriminatory validity and usefulness.<sup>84</sup>

The diagnostic specificity is of importance in treating the TCM pattern.<sup>9</sup> We found the same TCM patterns in healthy females/fertile and infertile women (paper III). How can it be understood that a specific TCM pattern, related to infertility, common in both fertile and infertile women, should explain infertility for some women? This raises the question of whether any pattern can really be specific to infertility. The TCM patterns are believed to reflect energetic imbalance, it is independent and different from a biomedical disease. Any energetic imbalance/TCM pattern can coexist with any biomedical disease. One may thus question the drive for connecting TCM patterns and biomedical diagnosis.

It is possible to diagnose healthy people as energetic imbalances may exist independent of experienced illness. Many of the symptoms and signs used to make TCM patterns do not necessarily indicate disease in a biomedical understanding, and thus foundation for combining TCM based diagnostics and biomedical diagnostics may be weak. For instance, in the interviews, all but one woman confirmed feeling stressed, as reported in Paper I. The participants were either first-time mothers or infertile, and both were considered in stressful situations. Moreover, stress could also be a common response to the current hectic society and indicate a general condition. Stress and emotion-related disorders are frequently associated with Liver Qi Stagnation, according to Scheid and Karchmer.<sup>27;90</sup> Stress can cause Qi Stagnation and depress the Liver Qi.<sup>16;30</sup> Liver Qi Stagnation causes overall obstruction and stagnation in the Uterus, among others.<sup>16;30</sup> Qi Stagnation in the Uterus might correspond to tubal obstruction,<sup>17</sup> and finally, the stagnation can lead to infertility. Consequently, according to this theory, stress can cause tubal obstruction and infertility, but it does not explain who are specifically affected. Not all who experience stress become infertile. The TCM lacks methods to determine whether a TCM pattern is related to stress and/or tubal obstruction, as well as to consider the degree of morbidity of the TCM pattern. According to TCM theories, irrespective of the missing link, it is important to treat the TCM pattern to prevent the development of disease. The treatment principle in TCM is to treat the TCM pattern.<sup>16;30</sup>

### **Treatment recommendations (Papers I and III)**

The acupuncture points discussed in Paper I and III were selected according to the TCM patterns. Since the inter-rater agreement on the TCM diagnoses was fair, the poor to moderate level of agreement on the recommended acupuncture points was expected. The



agreement improved when the acupuncture points were merged with their respective meridians and examined for a given merged TCM pattern. This showed that the two acupuncturists reached a better agreement on treating a defined TCM pattern. Zhang et al. also found that practitioners showed a better agreement on a diagnosis combined with its required herbal prescription compared to their agreement on their respective prescriptions.<sup>67</sup> This suggests that the acupuncturists and practitioners tried to practice the TCM principle in directing the treatment according to the pattern.

In the present study, the acupuncturists differed to a large extent in the frequency of their use of acupuncture points (Papers I and III). The explanation of the point selection was not reported. Sherman et al. asked acupuncturists to state the factors influencing their decision to prescribe specific acupuncture points for patients with chronic low back pain. They found that local pain points, meridian diagnosis and experience points that worked well regardless of the TCM diagnosis were used in 86%, 80% and 58% of the cases, respectively. The TCM theories were used less than 50% of the time.<sup>91</sup> Smith et al. used essential points, which were specified points applied during IVF treatments.<sup>92</sup> Likewise, di Villahermosa et al. reported a standard list of the most often recommended acupuncture points, according to the TCM patterns in infertile women.<sup>93</sup> Acupuncture points directed to a biomedical disease are termed formula points,<sup>1</sup> which are used in a westernised medical application of acupuncture. Formula points are criticised in practice since these are not based on TCM theories, which claim that individual treatment according to the TCM patterns provides the best outcome.<sup>9;34</sup> Borud et al, found that the acupuncturists had applied eight “core” points on all patterns related to menopause. They assumed that the points were used for symptomatic treatment rather than addressing the TCM pattern.<sup>13</sup> Some of these common acupuncture points; SP6, LR3, CV4<sup>13</sup> were also used to treat fertile and infertile in paper III. Textbook descriptions of the acupuncture points address both the TCM pattern as well as the relevant symptoms.<sup>94</sup> The mentioned points have widespread functions and they are for instance used for intestinal complains as diarrhoea, nausea, etc. and gynaecological symptoms as irregular menstruation, amenorrhoea, and uterine problems. They are used for palpitations, insomnia, dizziness, tinnitus and hypertension. Related to TCM patterns, SP6 is used among others to strengthen the Kidneys. LR3 spreads Liver Qi, subdues Liver Yang, and nourishes Liver Blood and Liver Yin. CV4 strengthen and nourishes the Kidneys,

benefits the uterus and assist conception. To strengthen the Kidneys can be relevant for infertile, fertile and menopause women, as illustrated by the functions of CV 4; weakness of lumbar region and legs, bones and joints in the middle aged and elderly, fear, fright, frequent urination, infertility, amenorrhoea, taxation heat, headache, and so on.<sup>94</sup> Hence, the acupuncture points can be used to treat a lot of physical and psychological conditions.

Cao et al. concluded in a meta-analysis that there was no significant difference between acupuncture treatment with or without TCM pattern differentiation in achieving better therapeutic effects.<sup>95</sup> The meta-analysis did not mention how the TCM patterns underwent quality assurance. According to TCM theories, a misdiagnosed TCM pattern should not provide a better outcome than treatment without TCM pattern differentiation. Taking into account the poor inter-rater reliability of TCM patterns and treatment recommendations, using formula points might be just as good. At least they secure a standard treatment that can be repeated, which is important for research.

The formula points used by Paulus et al. in treating infertility<sup>96</sup> have been replicated in other studies<sup>97</sup> and are also demanded in practice since they are regarded as having a positive effect on the pregnancy rate. Paulus et al. used a total of nine points, five before and four after an embryo transfer.<sup>96</sup> One of the points used before an embryo transfer was recommended in our study (Paper III).

The three most frequently used acupuncture points in the infertile group were CV 3, CV 4 and ST 29 (Paper III). These acupuncture points have all been recommended for infertility treatment.<sup>94</sup> Smith et al. listed all the three points as essential for acupuncture during an IVF cycle.<sup>92</sup> CV 3 and CV 4 were among the points that di Villahermosa et al. used to examine the influence of acupuncture outcomes on IVF.<sup>93</sup> This choice of points could therefore have been directed to infertility, not to the TCM pattern, since the two acupuncturists knew who of the women were infertile. The basis for point selection in Paper III was not examined and remained unknown. Two acupuncturists (as discussed in Paper II) asked for the biomedical diagnosis, implying that they considered the information useful for selecting acupuncture points, making the diagnosis and relating TCM patterns to fertility/infertility. It would seem that the acupuncturists used both formula points and acupuncture points, based on the

TCM patterns. This shows that there are several opportunities according to point selection, which can lead to poor inter-rater agreement on treatment recommendations.

Many practitioners advocate a standard treatment for each disease,<sup>20;77</sup> essential points<sup>92</sup> and formula points<sup>1</sup> are commonly used. Modern textbooks, as Maciocia's,<sup>17</sup> subdivides patterns according to biomedical disease categories, called type differentiation.<sup>20;77</sup> Type differentiations are theory based, and lack evidence for the coherence between TCM patterns and biomedical diseases. Deviations between authors on patterns related to a particular disease,<sup>8;17;30;31;33;37-50</sup> led to different approaches and treatment regimen depending on the textbook used.

Further, CM has been practised in various ways for thousands of years, a wide variety of diagnostic traditions coexisted in imperial China.<sup>20;77</sup> Low inter-rater reliability of TCM patterns seems to be a general problem, which means that people have been treated on different premises. The consequences of the myriads of acupuncture approaches, is that the TCM patterns seems to make no difference and that the treatment regimens become random. Hence, one may ask how important the TCM patterns are to guide treatment. Wang et al., claim that TCM patterns are products of speculation and imagery thinking, and the patterns are relatively generous, vague and abstractive.<sup>53</sup> TCM patterns are set on symptoms and signs that overlap across conditions. Therefore it might be favourable terms not to connect biomedical diseases and patterns. Hence, it is important for further research to explore the significance of the TCM patterns and TCM theories. Especially, this is required for acupuncture to maintain its place in academia and for acupuncturists to be accepted as health personnel.

## **Conclusion**

The low inter-rater agreement on TCM patterns indicates that the acupuncturists follow individual pattern differentiation processes. There are wide variations in how symptoms and signs are used to make diagnoses. Vague definitions of symptoms and signs in TCM theories might contribute to individual understanding and a personalised practice, leading to different TCM patterns for the same patients. Three of the eight TCM patterns associated with infertility (according to TCM theories) are confirmed, whereas two of the eight patterns are more frequently observed among the fertile women.

Our data suggest that the presence of some TCM patterns and the absence of others are associated with infertility. However, the sample size is small, and these data need more vigorous testing.

The poor agreement on diagnoses leads to the poor inter-rater reliability of treatment recommendations, posing challenges for individually tailored treatments and clinical trials in acupuncture.

Further research exploring the association between TCM patterns and biomedicine diseases, and also effect studies based on TCM patterns, should await better reliability on TCM diagnosis.

## Reference List

- (1) World Health Organisation Regional Office for the Western Pacific. World Health Organistaion International Standard Terminologies on Traditional Medicine in the Western Pacific Region. Geneva: WHO Library Cataloguing in Publication Data; 2007.
- (2) The European Information Center for Complementary and Alternativ Medicine (EICCAM) The European Information Centre for Complementary & Alternativ Medicine (EICCAM) 2012 Available from: URL:<http://www.eiccam.eu/pdfs/eiccambrochurecomplete.pdf>
- (3) Aarbakke, J. Norges offentlige utredninger; NOU 1998:21 Alternativ medisin. 1-405. Statens Forfaltningstjeneste Statens Trykning.
- (4) Fønnebø, V, Launsø, L, Bessesen, T., and Nilssen, T. Bruk av alternativ behandling i Norge. 1-12. Nifab.no. 26-8-2013.
- (5) NAFKAM NAFKAM-undersøkelsen 2012 27-10-2014 Available from: URL:[http://nifab.no/om\\_alternativ\\_behandling/tall\\_og\\_fakta/nafkam\\_undersoekelsen\\_2012](http://nifab.no/om_alternativ_behandling/tall_og_fakta/nafkam_undersoekelsen_2012)
- (6) Stener-Victorin E, Humaidan P. Use of acupuncture in female infertility and a summary of recent acupuncture studies related to embryo transfer. *Acupuncture in Medicine*. 2006;24:157-163.
- (7) Fertilitetssenteret Aleris sykehus Assistert Befruktning 2012 Available from: URL:<http://www.aleris.no/Global/Norge/Aleris%20Helse/Brosjyrer/IVF%20INFO%20BROSJYRE%20MAI%202012%201.pdf>
- (8) Maciocia G. Diagnosis in Chinese Medicine A Comprehensive guide. London: Churchill Livingstone; 2004.
- (9) Maciocia G. The Foundations of Chinese Medicine. A Comprehensive Text for Acupuncturists and Herbalists. London: Churchill Livingstone; 1989.
- (10) Liu T. Role of Acupuncturists in Acupuncture Treatment. *eCAM*. 2007;4:3-6.
- (11) Kaptchuk T. Chinese Medicine. The Web that has no Weaver. London: Rider; 1987.
- (12) Hu Zhen DF. Clinical Reasoning in Chinese Medicine. People's Medical Publishing House; 2008.
- (13) Borud, E. K., Alræk, T., White, A, and Grimsgaard, S. The acupuncture treatment for postmenopausal hot flushes (Acufash) study: traditional Chinese medicine diagnoses and acupuncture points used, and their relation to the treatment response. *Acupuncture in Medicine*. 2009; 27:101-108.
- (14) Sherman KJ, Cherkin DC, Hogeboom CJ. The Diagnosis and Treatment of Patients with Chronic Low-Back Pain by Traditional Chinese Medical Acupuncturists. *J Altern Complement Med*. 2001;7:641-650.
- (15) Maciocia G. Tongue Diagnosis in Chinese Medicine. Seattle: Eastland Press; 1987.

- (16) Maciocia G. *The Practice of Chinese Medicine. The Treatment of Diseases with Acupuncture and Chinese Herbs*. London: Chutchild Livinvstone; 1994.
- (17) Maciocia G. *Obstetrics & Gynecology in Chinese Medicine*. London: Churchill Livingstone; 1998.
- (18) Maciocia G. *The Channels of Acupuncture: Clinical Use of the Secondary Channels and Eight Extraordinary Vessels*. Elsevier Health Sciences; 2006. 1-768.
- (19) Taylor K. *Chinese Medicine in Early Communist China, 1945-63. A medicine of revolution*. London: Routledge Taylor & Francis Group; 2007.
- (20) Scheid V. *Chinese Medicine in Contemporary China, Plurality and Synthesis*. Durham & London: Duke University Press; 2002.
- (21) Gwei-Djen L, Needham J. *Celestial Lancets. A History and Rationale of Acupuncture and Moxa*. London: Routledge Curzon Taylor & Francis Group; 2002.
- (22) Veith I. *The Yellow Emperor's Classic of Internal Medicine. Huang Ti Nei Ching Su Wen*. London: University of California Press; 1972.
- (23) Gu P, Chen H. Modern bioinformatics meets traditional Chinese medicine. *Briefings in Bioinformatics*. 2013.
- (24) Cai J, Chao G, Chen D, Chen K, Chen X. *Advanced Textbook on Traditional Chinese Medicine and Pharmacology. Vol.1. First ed*. Beijing: New World Press; 1995.
- (25) Unschuld PU. *Medicine in China. A History of Ideas*. London: University of California Press; 1985.
- (26) Scheid, Volker. Convergent Lines of Descent: Symptoms, Patterns, Consellations, and the Emergent Interface of Systems Biology and Chinese Medicine. *East Asian Science, Technology and Society: An International Journal*. 2014; 8:107-139.
- (27) Karchmer, Eric I. The Excitations and Suppressions of the Times: Locating the Emotions in the Liver in Modern Chinese Medicine. *Cult Med Psychiatry*. 2013; 37:8-29.
- (28) Andrews, B. J. Wang Qingren and the History of Chinese Anatomy. *Journal of Chinese Medicine*. 1991; 30-36.
- (29) Karchmer El. Chinese Medicine in Action: On the Postcoloniality of Medical Practice in China. *Medical Anthropology*. 2010;29:226-252.
- (30) Flaws B. *A Handbook of Traditional Chinese Gynecology*. Boulder CO: Blue Poppy Press.; 1991.
- (31) Lyttleton J. *Treatment of Infertility with Chinese Medicine*. London: Churchill Livingstone; 2004.
- (32) Maciocia, Giovanni The view of tobacco in Chinese medicine 2014 Available from: URL:<http://maciociaonline.blogspot.no/2009/02/view-of-tobacco-in-chinese-medicine.html>

- (33) Wieseman N, Ye F. *A Practical Dictionary of Chinese Medicine*. Brookline, Massachusetts: Paradigm Publications; 1998.
- (34) Xinghua B. *Acupuncture: Visible Holism an original interpretation of acupuncture from Root to Tip*. Oxford: Butterworth Heinemann; 2001.
- (35) El-Toukhy T, Sunkara SK, Khairy M, Dyer R, Khalaf Y, Coomarasamy A. A systematic review and meta-analysis of acupuncture in in vitro fertilisation. *BJOG: An International Journal of Obstetrics & Gynaecology*. 2008;115:1203-1213.
- (36) Zegers-Hochschild, F., Adamson, G. D., Mouzon, J. de, Ishihara, O., Mansour, R. Nygren K., Sullivan, E., and Vanderpoel, S. International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary of ART terminology, 2009. *Fertility and sterility*. 2009; 92:1520-1524.
- (37) Farquhar, Judith. Objects, Processes, and Female Infertility in Chinese Medicine. *Medical Anthropology Quarterly*. 1991; 5:370-399.
- (38) Flaws B. *Endometriosis, Infertility And Traditional Chinese Medicine. A Laywoman's Guide*. Boulder, CO: Blue Poppy Press; 1989.
- (39) Flaws B. *My Sister The Moon. The Diagnosis & Treatment of Menstrual Diseases by Traditional Chinese Medicine*. Boulder, CO: Blue Poppy Press; 1992.
- (40) Flaws B. *A Handbook of Menstrual Diseases in Chinese Medicine*. Boulder, CO: Blue Poppy Press; 1997.
- (41) Guicheng X, Huiong G, Yuehua G, Yanhua S, Peigen L. *Concise Traditional Chinese Gynecology*. 1 ed. Jiangsu Science and Technology Publishing House; 1988.
- (42) Hou Jinglun. *Acupuncture and Moxibustion Therapy in Gynecology and Obstetrics*. Beijing, China: Beijing Science & Technology Press; 1995.
- (43) Lewis R. *The Ancient Chinese Wellness Program for Getting Pregnant and Having Healthy Babies. The Infertility Cure*. New York: Little, Brown and Company.; 2004.
- (44) Liang L. *Acupuncture & IVF. Increase IVF Success By 40-60%*. Boulder, CO: Blue Poppy Press; 2003.
- (45) Liu F. *The Essence of Liu Feng-wu's Gynecology*. Boulder, Co: Blue Poppy Press; 1998.
- (46) Low R. *A Thorsons Complementary Medical Textbook. Acupuncture in Gyneaecology and Obstetrics. An essential guide for practitioners and students*. London: Thorsons, Harper Collins Publishers; 1990.
- (47) Weixin J. *Diagnosis of Sterility and its Traditional Chinese Medicine Treatment*. Jinan, China: Shandong Science and Technology Press; 1999.
- (48) Yan Wu. *Practical Therapeutics of Traditional Chinese Medicine*. 1-716. 1997. Brookline, Massachusetts, Paradigm publications.
- (49) Yong T. *Gynecology of Traditional Chinese Medicine*. Shanghai: Publishing House of Shanghai University of Traditional Chinese Medicine.; 2002.

- (50) Zhang Ting-Liang. *A Handbook of Traditional Chinese Gynecology*. Ohio: Blue Poppy Press; 2005.
- (51) Fan KW. Pavlovian Theory and the Scientification of Acupuncture in 1950s China. In: Pei-kai Cheng, Fan KW, editors. *New Perspectives on the Research of Chinese Culture*. Singapore: Springer Singapore; 2013. 137-145.
- (52) Madaschi, C, Braga, DP, Figueira, Rde C, Laconelli, A Jr, and Borges, E Jr. Effect of acupuncture on assisted reproduction treatment outcomes. *Acupuncture in Medicine*. 2010; 28:180-184.
- (53) Wang, Jie, Wang, Pengqian, and Xiong, Xingjiang. Current Situation and Re-understanding of Syndrome and Formula Syndrome in Chinese Medicine. *Internal Medicine*. 2012; 2:1-5.
- (54) World Health Organization. Gender and Genetics. Assisted Reproductive Technologies (ARTs). <http://www.who.int/genomics/gender/en/index6.html> [ 2010 [cited 2015 Aug. 26]; Available from: URL:<http://www.who.int/genomics/gender/en/index6.html>
- (55) Altman DG. *Practical Statistics for Medical Research*. London: Chapman & Hall/CRC; 1999.
- (56) Crewson PE. Reader Agreement Studies. *American Journal of Roentgenology*. 2005;184:1391-1397.
- (57) Sim J, Wright CC. The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. *Physical Therapy*. 2005;85:257-268.
- (58) Viera, Anthony J and Garrett, Joanne M. Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine*. 2005; 37:360-363.
- (59) Andrew Mackinnon. DAG\_Stat. [http://www.mhri.edu.au/biostats/DAG\\_Stat/](http://www.mhri.edu.au/biostats/DAG_Stat/) [ 2000 [cited 2010 Jan. 1]; Available from: URL:[http://www.mhri.edu.au/biostats/DAG\\_Stat/](http://www.mhri.edu.au/biostats/DAG_Stat/)
- (60) Lydersen S, Fagerland MW, Laake P. Recommended tests for association in 2x2 tables. *Statist Med*. 2009;28:1159-1175.
- (61) Wiseman N, Ye F, Translation. Chinese medical pulse terms, taking account of the historical dimension. *Class Acup Orient Med*. 1999;1:55.
- (62) Zhang GG, Bausell B, Lao L. The variability of TCM pattern diagnosis and herbal prescription on rheumatoid arthritis patients. *Altern Ther Health Med*. 2004;10:58-63.
- (63) O'Brien KA, Birch S. A Review of the Reliability of Traditional East Asian Medicine Diagnoses. *The journal of alternative and complementary medicine*. 2009;15:353-366.
- (64) Mist S, Ritenbaugh C, Aickin M. Effects of Questionnaire-Based Diagnosis and Training on Inter-Rater Reliability Among Practitioners of Traditional Chinese Medicine. *J Altern Complement Med*. 2009;15:703-709.
- (65) O'Brien KA, Abbas E, Zhang J, et al. An Investigation into the Reliability of Chinese Medicine Diagnosis According to Eight Guiding Principles and Zang-Fu Theory in Australians with Hypercholesterolaemia. *J Altern Complement Med*. 2009;15:259-266.



- (66) Zhang GG, Singh B, Lee WL. Improvement of agreement in TCM diagnosis among TCM practitioners for persons with the conventional diagnosis of rheumatoid arthritis: Effect of training. *J Altern Complement Med.* 2008;14:381.
- (67) Zhang G, Lee W, Bausell B, et al. Variability in the Traditional Chinese Medicine (TCM) Diagnoses and Herbal Prescriptions Provided by Three TCM Practitioners for 40 Patients with Rheumatoid Arthritis. *J Altern Complement Med.* 2005;11:415-421.
- (68) Hogeboom CJ, Sherman KJ, Cherkin DC. Variation in diagnosis and treatment of chronic low back pain by traditional Chinese medicine acupuncturists. *Complement Ther Med.* 2001;9:154-166.
- (69) MacPherson H. Acupuncture for low back pain: traditional diagnosis and treatment of 148 patients in a clinical trial. *Complementary therapies in medicine.* 2004;12:38-44.
- (70) Lee, Ju Ah, Ko, Mi Mi, Kang, Byoung-kab, Alraek, Terje, Birch, Stephen, and Lee, Myeong Soo. Interobserver Reliability of Four Diagnostic Methods Using Traditional Korean Medicine for Stroke Patients. *Evidence-based Complementary and Alternative Medicine.* 2014; 2014:1-7.
- (71) Hua B, Abbas, Hayes A, et al. Reliability of Chinese Medicine Diagnostic Variables in the Examination of Patients with Osteoarthritis of the Knee. *J Altern Complement Med.* 2012;18:1-10.
- (72) Tang W, Lam M, Leung W, Sun W, Chan T, Ungvari GS. Traditional Chinese Medicine diagnoses in persons with ketamine abuse. *Journal of traditional Chinese medicine.* 2013;33:164-169.
- (73) Grant, Suzanne, Schnyer, Rosa, Hsu-Tung Chang, Dennis, Fahey, Paul, and Bensoussan, Alan. Interrater Reliability of Chinese Medicine Diagnosis in People with Prediabetes. *Evidence-based Complementary and Alternative Medicine.* 2013; 2013:1-8.
- (74) Kim M, Cobbin D, Zaslowski C. Traditional Chinese medicine tongue inspection: An examination of the inter- and intrapractitioner reliability for specific tongue characteristics. *J Altern Complement Med.* 2008;14:527-536.
- (75) King E, Cobbin D, Walsh S, Ryan D. The reliable measurement of radial pulse characteristics. *Acupuncture Med.* 2002;20:150.
- (76) Sung JY, Leung WK, Ching JYL. Agreements among traditional Chinese medicine practitioners in the diagnosis and treatment of irritable bowel syndrome. *Aliment Pharmacol Ther.* 2004;20:1205.
- (77) Shea, Jeannne L. Applying Evidence-Based Medicine to Traditional Chinese Medicine: Debate and Strategy. *The journal of alternative and complementary medicine.* 2006; 12:255-263.
- (78) Coeytaux RR, Chen W, Lindemuth CE, et al. Variability in the diagnosis and point selection for persons with frequent headache by traditional Chinese Medicine Acupunctuists. *J Altern Complement Med.* 2006;12:863-872.
- (79) O'Brien KA, Abbas E, Zhang J, et al. Understanding the Reliability of Diagnostic Variables in a Chinese Medicine Examination. *J Altern Complement Med.* 2009;15:727-734.

- (80) Su, Shi-Bing. Recent Advances in ZHENG Differentiation Research in Traditional Chinese Medicine. *International Journal of Integrative Medicine*. 2012; 1:1-10.
- (81) Fargen KM, Friedman WA. The Science of Medical Decision Making: Neurosurgery, Errors, and Personal Cognitive Strategies for Improving Quality of Care. *World Neurosurgery*. 2014;82:e21-e29.
- (82) Zhang GG, Bausell B, Lao L, Handwerger B, Berman BM. Assessing the consistency of traditional Chinese medical diagnosis: an integrative approach. 9, 66-71. 2003. *Altern Ther Health Med*.
- (83) de Sà Ferreira A. Advances in Chinese Medicine Diagnosis: From Traditional Methods to Computational Models. In: Haixue Kuang, editor. *Recent Advances in Theories and Practice of Chinese Medicine*. 12 A.D. 137-158.
- (84) Skjelstad, Dag Vegard. Er tiden moden for dimensjonale diagnoser? *Psykologtidsskriftet*. 2013; 50:552-558.
- (85) Andersson N, Klang B, Petersson G. Differences in clinical reasoning among nurses working in highly specialised paediatric care. *Journal of Clinical Nursing*. 2012;21:870-879.
- (86) Benner PE. *From Novice to Expert. Excellence and Power in Clinical Nursing Practice*. Commemorative ed. London: Alexander, Julie; Prentice-Hall International; 2001.
- (87) Rampp T, Tan L, Zhang L, Sun Zj, Klose P, Musial F et al. Menopause in German and Chinese women – An analysis of symptoms, TCM-diagnosis and hormone status.
- (88) Scheid V, Ward T, Tuffrey V. Comparing TCM textbook descriptions of menopausal syndrome with the lived experience of London women at midlife and the implications for Chinese medicine research. *Maturitas*. 2010;66:408-416.
- (89) Borud, E. K, Alraek, T., White, A., and Grimsgaard, S. The acupuncture on hot flashes among menopausal women (ACUFLASH) study: Observational follow up results at six and 12 months. *Menopause*. 2010; 17:262-268.
- (90) Scheid V. Depression, Constraint, and the Liver: (Dis)assembling the Treatment of Emotion-Related Disorders in Chinese Medicine. *Cult Med Psychiatry*. 2013;37:30-58.
- (91) Sherman KJ, Hogeboom CJ, Cherkin DC. How traditional Chinese medicine acupuncturists would diagnose and treat chronic low back pain: results of a survey of licensed acupuncturists in Washington State. *Complementary therapies in medicine*. 2001;9:146-153.
- (92) Smith C, Grant S, Lyttleton J, Cochrane S. Using a Delphi consensus process to develop an acupuncture treatment protocol by consensus for women undergoing Assisted Reproductive Technology (ART) treatment. *BMC complementary and alternative medicine*. 2012;12:88.
- (93) di Villahermosa DIM, dos Santos LG, Nogueira MB, Vilarino FL, Barbosa CP. Influence of acupuncture on the outcomes of in vitro fertilisation when embryo implantation has failed: a prospective randomised controlled clinical trial. *Acupuncture in Medicine*. 2013;31:157-161.

- (94) Deadman P, Al-Khafaji M. *A Manual of Acupuncture*. 1998. East Sussex, Journal of Chinese Medicine Publications.
- (95) Cao H, Bouchier S, Liu J. Does Syndrome Differentiation Matter? A Meta-Analysis of Randomized Controlled Trials in Cochrane Reviews of Acupuncture. *Medical Acupuncture*. 2012;24:68-76.
- (96) Paulus WE, Zhang M, Strehler E, El-Danasouri I, Sterzik K. Influence of acupuncture on the pregnancy rate in patients who undergo assisted reproduction therapy.[see comment]. *Fertility & Sterility* 77(4):721-4. 2002.
- (97) Domar, Alice, Meshay, Irene, Kelliher, Joseph, Alper, Michael, and Powers, Douglas. The impact of acupuncture on in vitro fertilization outcome. *Fertility and Sterility*. 2009; 91:723-726.
- (98) Alraek, Terje, Borud, Einar, and White, Adrian. Selecting Acupuncture Treatment for Hot Flashes: A Delphi Consensus Compared with a Clinical Trial. *The journal of alternative and complementary medicine*. 2011; 17:33-38.



