Complementary Therapies in Medicine (2014) 22, 954—964

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevierhealth.com/journals/ctim

Safety aspects of Chinese herbal medicine in pregnancy—Re-evaluation of experimental data of two animal studies and the clinical experience

Axel Wiebrecht[a,*], Wilhelm Gaus[b], Simon Becker[a], Josef Hummelsberger[a], Kirsten Kuhlmann[a]

[a] *Working Group ''Safety of Chinese Herbal Medicine During Pregnancy'' of the Schweizerische Berufsorganisation für TCM and the Societas Medicinae Sinensis, Bundesallee 141, D-12161 Berlin, Germany*
[b] *University of Ulm, Institute of Epidemiology and Medical Biometry, D-89081 Ulm, Germany*
Available online 19 August 2014

KEYWORDS
Chinese herbal medicine;
Pregnancy risks;
Embryotoxicity;
Fetotoxicity;
Teratogenicity;
Experimental studies;
Human data

Summary
Introduction: Chinese herbal medicine is an increasingly popular worldwide medical therapy which also has an impact in pregnancy. However, the question of its drug safety during pregnancy remains unresolved. Potential problems include teratogenicity, abortion, perinatal toxicity, pre- and postnatal developmental abnormalities, and eventually an increased risk for carcinomas in the offspring. Standard Materia Medica textbooks contain unreliable information when it comes to risks during pregnancy. Wang and co-workers conducted an experimental study (WS) on mice in which they investigated the effects of 17 Chinese medicinals regarding embryotoxicity and fetotoxicity. All these drugs seemed to exhibit multiple significant toxic effects. Another study by Li and co-workers (LS) investigated the reproductive toxicity of Atractylodis macrocephalae Rhizoma in mice, rats and rabbits. They described an increased pre- and postnatal mortality and, at high doses, congenital malformations. In an attempt to identify the risks of the tested medicinals during pregnancy, we analysed these two experimental studies and compared their results with possible safety data for humans from two reviews of clinical studies on threatened miscarriage (AR and CR).
Methods: We re-evaluated WS and LS in relation to accordance with internationally accepted rules, equivalence to human dose, biometric accuracy, plausibility, and coherence. Eligible studies of the two reviews on threatened miscarriage were evaluated for specific pregnancy risks concerning the 17 medicinals tested in WS and LS.
Results: We found that WS does not conform to international ICH guidelines and includes many inconsistencies, implausibilities and several severe biometrical flaws. It reported a total of 364 significant events out of which 145 false significant results are expected. The data-handling pointed to irregularities. Analysis of LS exhibited also many inconsistencies. The results

* Corresponding author. Tel.: +0049 30 8591067; fax: +0049 30 85999705.
E-mail address: axel.wiebrecht@tcm-embryotox.de (A. Wiebrecht).

http://dx.doi.org/10.1016/j.ctim.2014.08.005

regarding congenital malformations were statistically insignificant and are based on small case numbers. Insofar as the safety data of the 17 medicinals were documented by eligible studies of the two reviews, there was no indication of an increased abortion rate in humans. Fetal growth retardation was not observed in the human studies. For neonatal health and postnatal development, there were sufficient safety data only for a few medicinals in the human studies. As for teratogenicity, only small case numbers (0 to 109) were available from the human data.
Conclusion: WS and LS are not reliable data sources for deriving pregnancy risks in humans for the tested Chinese medicinals. In addition, the results appear to contradict the outcomes observed in the treatment of humans. Regarding teratogenicity, for most Chinese medicinals, neither the safety nor the risk during pregnancy can be definitively ascertained. Further studies on the risks of Chinese medicinals during pregnancy are urgently needed.

Contents

Introduction 955
Methods 956
Results 958
Re-evaluation of WS 958
Re-evaluation of LS 958
Re-evaluation of AR 958
Re-evaluation of CR 959
Discussion 959
Limitations 963
Conclusion 963
Conflict of interest statement 963
Acknowledgements 963
Appendix A. Supplementary data 963
References 963

Introduction

Chinese herbal medicine (CHM) is an increasingly popular medical therapy which is practiced world-wide.[1,2] Despite its ancient roots, Chinese medicine may be able to offer modern patients treatment options especially in cases in which Western medicine has not been able to provide satisfactory clinical results.[3] In order to fulfil its role as a treatment option for contemporary patients, CHM must conform to modern safety requirements. These are particularly crucial in the treatment of pregnant women because any therapy will affect the health of the developing embryo or foetus, respectively. Chinese medicine offers many treatment options for pregnancy-related indications such as threatened abortion, hyperemesis gravidarum, or intercurrent diseases. In treatment of patients for infertility, this may also impact the course of an undiagnosed pregnancy or any incipient pregnancy occurring after successful treatment.

The question of safety during pregnancy is already a difficult one for Western drugs due to insufficient data. An important concern is potential teratogenic risks. Other possible hazards include abortion, perinatal toxicity, pre- and postnatal developmental abnormalities, and an increased risk for carcinomas for the child later in his or her lifetime.

Standard Materia Medica textbooks providing excellent information about the properties, functions and actions of Chinese medicinals contain unreliable and sometimes even inconsistent information when it comes to risks during pregnancy. Examples of different safety classifications of some well-known standard textbooks[4–6] are provided in Table 1. In Chen and Chen[6] which focuses on Chinese pharmacology, terms such as embryotoxicity or fetotoxicity appear only twice (in relation to the plant substances Arecae Semen and Arecae Pericarpium).

In order to improve the unsatisfactory situation concerning the available data on pregnancy risks of CHM, Wang and co-workers conducted an experimental study (the ''Wang study'', WS) in mice in which they selected 17 Chinese medicinals commonly used during pregnancy and administered them at different periods of pregnancy.[7] Their effects regarding embryotoxicity and fetotoxicity were then investigated. The results caught the TCM community unaware. All the drugs investigated seemed to exhibit multiple significant toxic effects for several periods of drug administration, especially regarding fetal resorptions, stillbirths, fetal and postnatal deaths, postnatal growth retardation, and teratogenicity. Significant results for skeletal malformations were found for Rehmanniae Radix praeparata, Chuanxiong Rhizoma, and Citri reticulati Pericarpium. Minor malformations were found for Cuscutae Semen, Dipsaci Radix, Taxilli Herba, Glycyrrhizae Radix, Codonopsis Radix, Dioscoreae Radix, Amomi Fructus, Chuanxiong Rhizoma, Artemisiae argyi Folium and Citri reticulatae Pericarpium (Table 2). However, the results appear to show implausibilities and the effort to re-analyse the data was considered justified.

Table 1 Statements regarding safety during pregnancy in various standard textbooks.

Medicinal	pin yin	Bensky[4]	Chen[6]	Hempen[5]
Arecae Semen	*bing lang*	—	ca	ci
Coicis Semen	*yi yi ren*	—	ci	ca
Crataegi Fructus	*shan zha*	ci	—	al
Lycopi Herba	*ze lan*	ca	—	ci
Massa medicata ferm	*shen qu*	—	ca	ci
Plantaginis Semen	*che qian zi*	ca	—	ci
Siegesbeckiae Herba	*xi xian cao*	-	ca	ci

ci = ''contraindicated during pregnancy'', ca = ''use with caution during pregnancy'', al = ''allowed'', - = no indication of restrictions during pregnancy.

Another study investigated the reproductive toxicity of Atractylodis macrocephalae Rhizoma in mice, rats and rabbits (the ''Li study'', LS).[8] It used the same parameters and time frames as in WS. This study found that the administration of Atractylodis macrocephalae Rhizoma to mice was associated with significantly reduced fetal growth, a significantly prolonged duration of pregnancy and an increased prenatal and postnatal mortality. Furthermore, at high doses, congenital malformations (skeletal, fetal hydrops, and short ear anomaly) and fetal resorptions were described. The study was vividly documented with photos of the malformations in mice.

To explore the risks of using CHM during pregnancy in humans, a group of authors — with the participation of relevant authors from WS — conducted a review and meta-analysis of adverse outcomes in studies and case series on threatened miscarriage (the ''adverse outcome review'', AR).[9] According to the authors, threatened miscarriage is the most common indication for CHM during pregnancy in China. Another evaluation of CHM in treating threatened miscarriage is the *Cochrane review* on this subject (CR).[10] This review included for the most part different studies than those of AR.

Although the methodological quality of most of the trials included in these two reviews does not permit definite conclusions to be drawn, the data from eligible studies can nonetheless be used to evaluate potential pregnancy risks of the used medicinals. Through an analysis of the extracted safety data on humans and comparison with the experimental data of the animal studies, this paper tries to add to the discussion on the safety of Chinese medicines during pregnancy.

Methods

WS was analyzed in relation to accordance with internationally accepted rules, equivalence to human dose, biometric accuracy, plausibility and coherence. Similarly, LS was assessed for equivalence of dosing, plausibility and coherence.

AR was evaluated for specific pregnancy risks concerning the 17 medicinals tested in WS and LS. From this review, we included those studies which implied a documentation of the respective pregnancy risk and ensured a sufficiently long observation time for this scope. The following pregnancy risks were evaluated and compared with the results of WS: abortion rate (in relation to fetal resorptions in WS), fetal growth retardation and birth weight (in relation to fetal growth in WS), neonatal health and development (in relation to postnatal deaths and postnatal weight gain in WS) and congenital malformations. To ensure a high validity, only studies were included which documented a sufficient follow-up time: for fetal growth retardation and birth weight the minimum time was until delivery, for congenital malformations, neonatal health and development until the fourth postnatal week.

We excluded studies which are obviously not credible because there were significantly fewer side-effects described as could be spontaneously expected, when external applications of CHM were investigated or if there was no information provided about the medicinals used or the number of patients treated with the different medicinals.

For each study relevant to a particular question, we extracted the cases in which a certain medicinal in the context of a formula was used. Subsequently, the case numbers of the relevant studies were added. The sum was the number of documented cases for which data with respect to a certain adverse effect were available. For each pregnancy risk and each particular medicinal, a number of at least 100 documented cases demonstrating no increased rate of adverse events was rated as a preliminary evidence of safety. With regard to teratogenicity, however, a sample size of 300 was considered to be the minimum. The spontaneous rate for major malformations lies at 3 to 4 percent for all newborns.[11] Three hundred documented cases allow to exclude a 2.7-fold increase of this risk (at a power of 80 percent, 95 percent confidence level).[12] In our analysis, we have listed only drugs that have been tested in WS or LS, respectively. In the clinical studies, most formulas included further drugs that we did not analyze.

The same methodology was applied to eligible studies from CR. For this analysis only the abortion rate could be evaluated because other adverse events are not adequately documented. For those controlled trials that in the meta-analyses of AR or CR had shown a significant superiority of the combined therapy compared with Western drugs alone, the numbers of cases for each medicinal were summed-up. The sum may indicate the number of documented cases in which a lack of abortifacient effect of the medicinal in question is supported.

Table 2 ''Significant'' results in WS[7] regarding reproductive toxicology for different periods of pregnancy at which the medicinals were given.

Tested for period:		FR	CRL	HL	SO	SkA	MiA	SB	EPD	LPD	PWG	PWG28
		a–b–c	b–c	b–c	b	c–d–e	c	d–e	d–e	d–e	d–e	d–e
Medicinal	English name											
Atractylodis macrocephalae Rhizoma	Largehead Atractylodis Rhizome	a,c	b,c	b,c	b			d	d,e	e	d,e	
Cuscutae Semen	Chinese Dodder Seed	b,c	b,c		b		c		e		d,e	d
Dipsaci Radix	Himalayan Teasel Root	b,c	b,c		b		c	e	d,e		d	d
Taxilli Herba	Chinese Taxillus Twing	a,b,c	b,c				c	e		d	d,e	d
Glycyrrhizae Radix	Liquorice Root	a	b,c	b,c	b		c		e		d,e	
Astragali Radix	Milkvetch Root	a,c	b	b	b			e	e		d,e	
Paeoniae Radix alba	White Paeony Root	a	b,c					e	e	d	d,e	e
Angelicae sinensis Radix	Chinese Angelica	a,b,c	b,c	c	b			d	e	d	d,e	
Eucommiae Cortex	Eucommia Bark	b	b,c	b,c	b			e			d,e	
Rehmanniae Radix praep.	Steamed Rehmannia Root	c	b,c	c	b	c			e		d,e	d
Codonopsis Radix	Pilose Asiabell Root	a,b	c	c			c		d		d	
Dioscoreae Radix	Common Yam Rhizome		b,c	b,c			c		d	d	d,e	d,e
Amomi Fructus	Villous Amomum Fruit	a,b	b,c	c			c		e	d	d,e	
Rehmanniae Radix	Rehmannia Root	c	b					e	d,e	d,e	d,e	d,e
Chuanxiong Rhizoma	Szechuan Lovage Rhizome	a,b,c	b,c	b,c		d	c	e	e	d	d,e	
Artemisiae argyi Folium	Chinese Mugwort Leaf	a,b	b,c	c			c		e		d	
Citri reticulatae Pericarpium	Tangerine Peel	c	b	c		c,d	c			d	d,e	

Period (a) gestational day (GD) 3–6, period (b) GD 6–8, period (c) GD 8–15, period (d) GD 15 until delivery, period (e) GD 0 until delivery.
FR = fetal resorptions, CRL = crown-rump length, HL = head length, SO = somite, SkA = skeletal anomalies, MiA = other minor anomalies, SB = stillbirth, EPD = early postnatal death, LPD = late postnatal death, PGW = postnatal weight gain, significant results for at least one period (day 1, 7, 14, 28), PWG28 = postnatal weight gain, significant results at postnatal day 28.

Results

Re-evaluation of WS

In our analysis, we found that WS does not conform to the international ICH guideline[13] which specifies the rat to be the preferred animal species in testing reproductive toxicity. In addition, the usual number of cases in studies to test for developmental toxicity is set at 20 dams,[14] but in WS the number of dams ranged from 4 to 10. In this respect, the probability of random errors is high.

The authors claimed that they maintained the dosage in a range equivalent to the clinical dose for humans and a 2- or 3-fold increase thereof, respectively. However, the basic study dose slightly exceeded the maximum of the equivalent dose listed in the Chinese Pharmacopeia for four drugs. For another five drugs (Cuscutae Semen, Taxilli Herba, Eucommiae Cortex, Rehmanniae Radix praep. and unprepared Rehmanniae Radix), the dose level clearly exceeded the maximum.[15]

A closer look at WS reveals many inconsistencies. For medicinals showing toxic effects at certain periods of administration during pregnancy, the toxicity should be particularly evident if the medicinals were taken throughout the entire pregnancy. However, skeletal anomalies reported for three medicinals (Rehmanniae Radix praep., Chuanxiong Rhizoma, and Citri reticulatae Pericarpium) given in certain periods did not arise when these medicinals were administered throughout the entire pregnancy. The same is true for two medicinals (Atractylodis macrocephalae Rhizoma and Angelicae sinensis Radix) associated with stillbirth, for two medicinals associated with early postnatal death and for seven medicinals associated with late postnatal death (see Table 2). It seems implausible that the toxic effects appearing at one period of pregnancy can disappear after prolonged administration of the medicinal.

WS appears to contain several serious biometrical mistakes. A total of 1751 animals were treated with Chinese medicinals (study group) but the control group consisted of merely 46 animals. It is very likely that in a study group that is approximately 38 times larger than the control group, more abnormalities should arise. Even if the rates for toxic effects were the same in the study and control groups, due to the very large difference in number of cases, about 38 times more events can be expected for the study group.

WS reported 364 significant events. Of these, 211 had a significance level of 5%, 47 had a significance level of 1% and 106 a significance level of 0.1%. The large number of significant events at 0.1% is surprising because at this level of significance only 1/10 of the significant events from a 1% significance level is to be expected.

In Tables S3—S7 of WS, 17 substances, 3 doses and (total of all tables) 57 criteria are listed. Thus, $17 \times 3 \times 57 = 2907$ tests are possible. We presume that all these tests were calculated. But even if only a portion of these tests had been calculated, it may be assumed that the most ''promising'' tests were included. Simply by applying the significance level of 5%, a total of 145 ($2.907 \times 5\%$) false significant results of the total of 364 significant events are expected. Unfortunately, Wang et al. did not mention this problem of multiple statistical tests.

For the statistical analysis, the authors have used the data from every control group for 51 tests. The multiple use of the same control group requires an adjustment of the *p*-values. The authors did not consider this.

In addition to the biometric deficiencies, the handling of the data appears to be incorrect. All values for Atractylodis macrocephalae Rhizoma, including the means and standard deviations are identical to those of LS. The only difference is that the number of cases was doubled; the number of dams was doubled from 5 to 10 and from 4 to 8, respectively. The same was done for the numbers of offspring. A coincidence can thus be ruled out. Nor is it likely that the numbers in LS were halved. Doubling the number of cases and non-disclosure of using results from another study is not a reputable research approach. The approach of this study raises serious doubts about the reliability of its findings.

Re-evaluation of LS

In the case of Atractylodis macrocephalae Rhizoma, the maximum dose of 12 g for humans[15] was used as a baseline and not the median dose as claimed. Therefore, the 1- to 3- (mice, rabbits) or 6-fold (rats) equivalent dose, respectively was used in the experimental study.

When the results with mice are analyzed, a general dose-response relationship cannot be detected, even if only significant findings are considered. The crown-rump length was reduced if Atractylodis macrocephalae Rhizoma was administered during the period of gastrulation and increased when administered in the organogenesis period. But the latter held true only for a 1-fold clinical dose, not at higher doses. If Atractylodis macrocephalae Rhizoma has toxic effects, the administration during the *entire* pregnancy period should permit a conclusion. However, there are 13 cases of late postnatal death with the 2-fold dose, but no case with the 3-fold dose. Only one single significant result is noted for the highest dose: a reduced weight gain at postnatal day 7, but not at postnatal day 14, 21 or 28. There were no significant differences for other parameters such as stillbirths, postnatal deaths and deformities compared to control.

These results are inconsistent and may be due merely to chance, perhaps caused by the small case number of 5 dams which is clearly below the customary number of 20. In rats and rabbits also, there were no significant differences when the 3- or 6-fold ''clinical dose'' was administered. In this case, only 3 dams with their offspring were tested.

The results regarding congenital malformations were statistically insignificant. Results that were not significant and are based on such small numbers of cases cannot establish a causal relationship regarding teratogenicity.

In summary, no valid conclusions with respect to embryotoxic or fetotoxic effects for Atractylodis macrocephalae Rhizoma in mice, rats and rabbits can be drawn from this study.

Re-evaluation of AR

The results of the experimental studies WS and LS in animals can be contrasted with the human experiences of adverse events in the treatment of threatened abortion. If

the *increased resorption rates* of fetuses found in WS for all medicinals except Dioscoreae Rhizoma were transferable to humans, this would mean that the administration of these medicinals to pregnant women, even more a combination of those, would have inevitably led to serial abortions or stillbirths, respectively. In AR the abortion rate of controlled trials with CHM alone was at 5 to 18.5% and in the case studies at 2 to 20%, with combined therapy at 0 to 22.3%, and with Western drugs alone at 15 to 33%. In this respect the results are consistent and there is no evidence of an increased risk of abortion from Chinese medicinals.

For determining the frequency at which the individual medicinals are administrated to humans we considered controlled trials and case control studies that met our inclusion criteria. The study by Zhang (2000)* was not included in this and the following analyses, even though it contains the highest number of cases ($n = 630$). It was determined that this study lacks sensitivity for adverse events as it documents considerably fewer adverse events than can be spontaneously expected in uncomplicated pregnancy. The number of cases treated for threatened miscarriage and documenting safety regarding an increased rate of abortion is shown in Table 3.

For 13 of the total 17 medicinals tested in WS and LS, respectively, in mice, documented experiences on humans with case numbers of at least 100 are available (range of 100 to 1502 cases). For one medicinal (Chuanxiong Rhizoma), there are only 44 documented cases; for 3 more medicinals, no cases exist because these drugs are not indicated for treatment of threatened abortion (Citri reticulatae Pericarpium, Artemisiae argyi Folium, and raw Rehmanniae Radix). For 11 medicinals, case numbers of more than 500 exist, for six medicinals more than 1000. Thus, for most of the medicinals tested in WS and LS there is quite a high number of cases documenting the safety with respect to abortion.

Growth retardation (in terms of crown-rump length, head length, somite) is present in WS for all tested medicinals for one or two of the testing periods. For documented medicinals no evidence of fetal growth retardation was observed in humans. For 11 medicinals, the number of cases exceeded 100 (Atractylodis macrocephalae Rhizoma, Cuscutae Semen, Dipsaci Radix, Taxilli Herba, Glycyrrhizae Radix, Astragali Radix, Paeoniae Radix alba, Angelicae sinensis Radix, Eucommiae Cortex, Dioscoreae Rhizoma and Amomi Fructus). For six medicinals, smaller case numbers or no results are available (Table 4 and Supplementary Table S3).

WS found increased rates of early or late *postnatal death* in most of the tested medicinals. The only exceptions were Taxilli Herba, Eucommiae Cortex and Citri reticulatae Pericarpium. For the comparison with experiences on humans, the parameters were neonatal health and development. Three studies of AR fulfilled the requirements defined for this evaluation. Relevant case numbers of more than 100 documented cases were reached only for Atractylodis macrocephalae Rhizoma, Dipsaci Radix and Taxilli Herba (Table 4 and Supplementary Table S4).

With respect to *teratogenicity*, among the studies of AR only one unspecified malformation was reported in the case series Zhou (2006)*. With a sample size of 40 pregnant women treated with Chinese medicine, this case (corresponding to 2.5%) is not significant. Because the study's observation period ends at delivery, it does not meet our inclusion requirements. For individual medicinals, only small numbers of cases (0 to 109) are available (Table 4 and Supplementary Table S5). These are insufficient for a safety assessment regarding congenital malformations. CHM was administered in the study conducted by He (1997)* between the 6th and the 12th week of pregnancy, and in the study conducted by Wang and Li (2000)* between the 5th and 20th week. This partly exceeds the sensitive period for malformations.

For all tested medicinals, WS found a ''significant'' impairment of *postnatal weight gain* on at least one reference date. Taking into consideration only the last studied reference date (postnatal day 28), there were ''significant'' weight deficits noted for seven medicinals. In humans, however, the three studies with an observation period of one year or more listed in Table 4 and S4 showed no postnatal developmental disorders. Sufficient sample sizes of over 100 are documented for Atractylodis macrocephalae Rhizoma, Dipsaci Radix and Taxilli Herba.

Re-evaluation of CR

In CR a meta-analysis was undertaken which included five randomized controlled trials. The combination therapy of CHM and Western drugs versus Western drugs alone was significantly superior in terms of the main outcome variable ''absence of abortion''. Of these five studies, one provided no information about the administered medicinals and another used an external application. We calculated the case numbers of the remaining three studies. As to Zhong's (2002)* study, only the case number for the combined therapy was included.

The resulting case numbers for individual medicinals were pooled with the case numbers of the AR meta-analysis, as well for combined therapy of CHM and Western drugs. Two studies of AR were not included due to an insufficient follow-up period. Table 5 shows the number of documented cases for medicinals for which a significantly reduced rate of miscarriage was shown as part of a combined therapy versus Western drugs alone. This is a much stricter criterion than applies to the case numbers in Table 3. These results can be used to argue against the transferability of a potential fetal die-off in mice to humans. Eleven drugs were present in case numbers of 100 or more, namely Atractylodis macrocephalae Rhizoma, Cuscutae Semen, Dipsaci Radix, Taxilli Herba, Glycyrrhizae Radix, Paeoniae Radix alba, Eucommiae Cortex, Rehmanniae Radix praeparata, Codonopsis Radix, Dioscoreae Rhizoma, and Amomi Fructus. Due to the lack of precise information for other medicinals, for these only minimum values that lie below 100 could be estimated.

Discussion

The authors of WS attempted to identify risks for the application of Chinese medicinals during pregnancy. The authors were interested in conducting a carefully planned study, for example they tried to use doses akin to those administered to humans.

However, their work contains serious biometric flaws such as in the scheduling and evaluation of the study and

Table 3 Number of cases on which single medicinals were used in relevant studies of AR[9] showing no evidence of failure of intervention at threatened miscarriage.

Study	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	Σ
Number of cases CM	54	45	100	44	40	131	56	58	118	58	68	47	86	60	30	305	67	34	40	61	41	**1543**
Formula	ZXBTD	BSGTD	JWATD	TSPSP	STP	ZNBTF	YSGCD	TEAP	STP	STP	BYD	STP	STP	STP	STP	ATD	ATD	ATD	STP	STP	WZD	
Atraclyl. macr. Rhz.	54		100	44	40	131	56		118	58	68	47	86	60	30	305	67	34	40	61		**1399**
Cuscutae Sem.	54	45	100	44	40	131	56	58	118	58	68	47	86	60	30	305	67	34	40	61		**1502**
Dipsaci Rd.	54	45	100	44	40	131	56		118	58	68	47	86	60	30	305	67	34	40	61		**1444**
Taxilli Hb.	54	45		44	40	131	56		118	58	68	47	86	60	30	305	67	34	40	61		**1344**
Glycyrrhizae Rd.	54	45		44												305	67	34				**549**
Astragali Rd.	54			44		131	56	58								305	67	34				**749**
Paeoniae Rd. alba	54	45	100	44			56	58								305	67	34				**763**
Angelicae sinen. Rd.	54			44				58								305	67	34				**562**
Eucommiae Cort.		45	100	44	40		56		118	58		47	86	60	30	305	67	34	40	61	41	**1232**
Rehmanniae Rd. prp.				44			56															**100**
Codonopsis Rd.		45	100	44		131					68											**388**
Dioscoreae Rhz.		45			40		56		118	58		47	86	60	30	305	67	34	40	61		**1047**
Amomi Fr.	54		100		40				118	58		47	86	60	30				40	61		**694**
Chuanxiong Rhz.				44																		**44**

Only studies were included that exhibited a sufficient long follow-up period. Only medicinals are presented that are tested in WS. Studies: (1) Song and Zhu (2007), (2) Li (2006), (3) He and Che (2007), (4) Yue (2009), (5) Zhou (2006), (6) Xu (2008), (7) Ye and Qiu (2008), (8) Luo (2007), (9) Chou (2002), (10) Xu (2001), (11) Chen and Yun (1999), (12) Cui (1998), (13) Kang (1998), (14) Chen (1997), (15) He (1997), (16) Zhou (1997), (17) Zhu and Li (1992), (18) Tian and Li (1991), (19) Li (1989), (20) Wang and Wang (1987), (21) Wu (1987). Formulas: ATD—An Tai Decoction, BSGTD—Bu Shen Gu Tai Dec., BYD—Bao Yun Dec., JWATD—Jiu Wei An Tai Dec., STP—Shou Tai Pill, TEAP—Tai Er An Pill, TSPSP—Tai Shan Pan Shi Pill, WZD—Wu Zi Dec., YSGCD—Yi Shen Gu Chong Dec., ZNBTF—Zi Ni Bao Tai Formula, ZXBTD—Zhi Xue Bao Tai Dec. CM = Chinese herbal therapy (CHT) alone or combined medicine (CHT plus Western drugs). References of the studies are quoted in Supplementary Table S1.

Table 4 Sum of frequencies the various medicinals were used in relevant studies of AR[9] apparently documenting safety with regard to the specific outcome.

Medicinal	No increased rate of abortion	Fetal growth or birth weight	Neonat. health and develop.	Congenital malform. and neonat. develop.
Atraclyl. macr. Rhz.	1399	562	176	109
Cuscutae Sem.	1502	483	97	30
Dipsaci Rd.	1444	562	176	109
Taxilli Hb.	1344	613	176	109
Glycyrrhizae Rd.	549	121	67	0
Astragali Rd.	749	172	67	0
Paeoniae Rd. alba	763	172	67	0
Angelicae sinen. Rd.	562	121	67	0
Eucommiae Cort.	1232	470	97	30
Rehmanniae Rd. prp.	100	51	0	0
Codonopsis Rd.	388	0	0	0
Dioscoreae Rhz.	1047	429	97	30
Amomi Fr.	694	416	30	30
Rehmanniae Rd.	0	0	0	0
Artemisiae argyi Fol.	0	0	0	0
Chuanxiong Rhz.	44	0	0	0
Citri reticulatae Peric.	0	0	0	0

Only medicinals are presented that are tested in WS. The number of cases is summed up from the relevant studies.
Details are presented in Supplementary Tables S3–S5.

Table 5 Number of cases of studies showing a significant lower abortion rate for combined medicine (CHT plus Western drugs) versus Western drugs alone at threatened abortion.

Study	Li (2006)†	He and Che (2007)†	Yue (2009)†	Chen (2002)††	Lv (2007)††	Zhong (2002)††	Number of cases **Total**
Number of cases CM	**45**	**100**	**44**	**51**	**58**	**30**	**Total**
Atraclyl. macr. Rhz.		100	44	n.d.	58	30	**≥232**
Cuscutae Sem.	45	100	44	51	58	30	**328**
Dipsaci Rd.	45	100	44	51	58	30	**328**
Taxilli Hb.	45		44	51	58	30	**228**
Glycyrrhizae Rd.	45		44	n.d.	58		**≥147**
Astragali Rd.			44	n.d.			**≥44**
Paeoniae Rd. alba	45	100	44	n.d.	n.d.	n.d.	**≥189**
Angelicae sinen. Rd.			44	n.d.			**≥44**
Eucommiae Cort.	45	100	44		n.d.		**≥189**
Rehmanniae Rd. prp.			44		58		**102**
Codonopsis Rd.	45	100	44	n.d.	58	30	**≥277**
Dioscoreae Rhz.	45				58	30	**133**
Amomi Fr.		100		n.d.	n.d.	n.d.	**≥100**
Rehmanniae Rd.				n.d.		n.d.	**≥0**
Artemisiae argyi Fol.				n.d.		n.d.	**≥0**
Chuanxiong Rhz.			44				**44**
Citri reticulatae Peric.				n.d.		n.d.	**≥0**

''Number of cases total'' is the number of cases from relevant AR[9] and CR[10] studies summed up.
Studies exhibiting an insufficient follow-up were excluded. Only medicinals are presented that are tested in WS and LS. Some studies use more medicinals in addition to the basic formula, whose frequency is not defined (''n.d.''), case numbers are minimum numbers here. CM = CHT alone or combined medicine (CHT plus Western drugs). References of the studies are quoted in Supplementary Tables S1 and S2.
† Studies from the AR meta-analysis.
†† Studies from the CR meta-analysis.

irregularities in handling of data. The number of cases is small. These factors prevent us from accepting the results as valid. The absence of a dose-response relationship makes the results appear even more implausible. Thus, we conclude that the study is not suitable for using the results to derive pregnancy risks in humans for the tested Chinese medicinals. Similarly, because of several implausibilities and insignificant results, LS cannot be considered qualified for drawing valid conclusions with respect to embryotoxicity or fetotoxicity of Atractylodis macrocephalae.

Several results of WS and LS differ from the outcomes of other research. WS found developmental disorders with the use of Taxilli Herba in mice embryos. Liu et al. tested this medicinal in groups of 12 pregnant rats and did not detect these disorders. There were neither divergences in terms of body weight, body length, or tail length of the embryos compared to control nor were there skeletal anomalies.[16] The conclusions of LS clash with results from another study for which a detailed abstract was published.[17] This study examined the pregnancy risks of Atractylodis macrocephalae Rhizoma in groups of 17 or 18 mice. An aqueous extract of Atractylodis macrocephalae Rhizoma was administered orally in doses of 2, 8 and 32 g/kg from day 6 to 15 of gestation. Distilled water served as the control. The study found no significant differences between the varying doses of Atractylodis macrocephalae Rhizoma and the control with respect to the incidence of fetal resorptions, dead fetuses, obvious malformations or skeletal abnormalities.

In addition, the results of WS and LS appear to contradict the results obtained in treatment of humans, in so far as the administration of the tested medicinals has been documented in Chinese controlled trials or case series and were included in AR.[9] For the parameters abortion risk, fetal growth retardation, postnatal survival and postnatal development, the number of at least 100 cases was rated as minimum to make a preliminary statement on the safety of Chinese medicinals. For the parameter abortion rate, most drugs are reaching high numbers of cases of up to more than 1500. Thus, administration of the medicinals in question is unlikely to increase the rate of abortion. This statement is further supported by studies from CR[10] on the treatment of threatened miscarriage. When combining the results of the two meta-analyses with the strict criterion of significant superiority of CHM as adjunctive therapy compared to Western drugs alone, eleven out of 17 drugs achieved a number of 100 documented cases or more. A further review on CHM for recurrent miscarriage evaluating studies different from those of CR showed that CHM or CHM combined with Western drugs was superior to Western drugs alone in most cases. There was no evidence of an elevated abortion rate by CHM.[18]

Fewer cases are available for analysis to answer the questions of fetal growth retardation, postnatal deaths and postnatal developmental disorders. However, the documented experiences in humans appear to disprove the occurrence of such effects. For teratogenicity, the number of cases was insufficient to ascertain the absence of this risk. Otherwise, it is worth noting that in AR there are no positive clues of teratogenicity for the medicinals tested.

Generally, the transferability of results from experimental animal studies to the therapeutic situation in humans is questionable.[11,19] Species-related bioavailability, metabolism, and sensitivity can differ greatly and — given the differing embryogenesis — can lead to different effects. Since there is a threshold dose for teratogenic effects, results from high doses often used in animal studies are not generally predictive for therapeutic doses of humans. There are several substances which have proven teratogenic effects in animals which do not exhibit this property in humans. For example, aspirin is known to cause cardiac defects in rats and rhesus monkeys but shows no teratogenicity in humans. Insulin is teratogenic in rats, mice and rabbits, but is regarded the treatment of choice for diabetes during pregnancy in humans.[14] Of 165 compounds which are regarded as non-teratogenic to humans, the tests were actually negative only in 80 percent for monkeys, 70 percent for rabbits, 50 percent for rats and in only 35 percent for hamsters and mice.[20] There is also evidence of teratogenic or embryotoxic and fetotoxic effects from animal studies for food constituents, e.g. for caffeine[20] and for alkaloids occurring in potatoes which however, pose no known risk to humans.[21–23]

For natural medicines with a long history of use, experimental animal studies are likely to raise more questions than provide answers. Instead, more data from human studies are needed. Intervention trials on the safety of drugs during pregnancy are forbidden for ethical reasons. Human data on teratogenicity are derived primarily from cohort studies with large numbers of cases in which a relationship between the occurrence of congenital malformations and the use of a certain drug can be established. These studies document the outcomes of drugs that were taken without or despite knowledge of an existing pregnancy and compare these data with the frequencies of malformations with those of non-exposed pregnancies.

At least one cohort study has shown a risk of congenital malformations for the use of CHM in humans.[24] Coptidis Rhizoma (*huang lian*) was associated with increased malformations of the nervous system and the external genital system. The formula *an tai yin* led to increased malformations of the musculoskeletal and connective tissue and the eye. (The composition of *an tai yin* varies between sources.) In children whose mothers had taken Coptidis Rhizoma during pregnancy, the same authors were able to observe an increase in cancer cases, particularly of the CNS, after a mean follow-up period of 14.9 years.[25]

Usually, the use of natural remedies during pregnancy is not documented. The reasons for this are, among others, that these remedies are considered less risky and that their consumption frequently occurs without medical supervision. As a result, little objective data about their teratogenicity and other pregnancy risks are available.[19,26–28] And in fact, even less data are available than for pharmaceutical drugs that are relatively new on the market.[11] Nevertheless, the experience collected over time is of some value, even in terms of teratogenicity, since CHM has been used under modern scientific supervision for about two generations. Even during this time, most teratogens of Western drugs were first noticed through case reports and clinical case studies and not by animal testing.[19]

While conventional medicine often takes the view that natural remedies are to be avoided during pregnancy until their safety has been proven,[29,30] others argue that traditionally-used remedies with few side-effects should not

be banned until there is clear evidence of risk.[26] In the case of dietary supplements, the European Food Safety Authoritiy (EFSA) has established a category in which a substance is presumed to be safe and does not require further safety assays if no adverse effects have been reported after long-term use on large populations[31]: ''Depending on the botanical ingredient and its uses, there are circumstances under which no additional data are judged necessary for the safety evaluation, i.e. a presumption of safety would be applied. This would be the case whenever available data would allow to conclude that exposure to known levels of the botanical ingredient has occurred in large population groups for many years without reported adverse effects.''

On the other hand, experience alone is not sufficient to detect seldom-occurring risks. The obvious teratogenicity of Coptidis Rhizoma was only recognized by a large epidemiological study. Even if this cannot be regarded as a definitive statement, this risk must be perceived and be closely observed as long as no mitigating data are available. Unfortunately, this study is an exception and there is a continuing lack of knowledge about the teratogenicity of Chinese medicinals. For most of conventional drugs, too, there is considerable uncertainty regarding their pregnancy risks, even for those that are sometimes administered during pregnancy. It is estimated that about 98 percent of the drugs approved between 2000 and 2010 in the U.S. fall into the category of ''undetermined teratogenic risk''.[32]

Therefore, as for chemical drugs, Chinese medicinals should only be administered to pregnant women after a careful benefit-risk assessment. Medicines should only be applied if there is an unequivocal indication, especially during the first trimester. The same applies for fertility treatment, which often extends into the pregnancy. The important period of organogenesis in humans is between gestational day 22 and 55. However, it is known that some teratogens can cause malformations if administered prior to this period.[14] For example, even one year after discontinuing the anti-psoriatic agent acitretin, every 20th child could still be affected by malformations.[33,34]

Limitations

One limitation of the risks assessment is the fact that the effects as well as the side-effects of medicinals may depend on the type of combination in which they are administered. However, to assess each individual combination by itself would require a great deal more data. This type of data is generally unavailable. Furthermore, in some studies, the standard formula was modified by adding or subtracting medicinals individually, a common practice in Chinese herbal medicine. At present, any evaluation of commonly-used medicinals is only possible by relying on the sum of experiences collected through the application in different combinations.

The data extracted from AR and CR are based on low-quality studies. Therefore, only limited conclusions regarding both efficacy and safety can be made from these. Furthermore, the Chinese medicinals were used at different times during pregnancy, in some cases outside the sensitive phase for malformations. For a few studies no data for the time of administration are available. On the other hand, some data are based on a high number of cases and are very consistent, which partly outweighs the limitations.

Conclusion

Taking the viewpoint that documented clinical experience is of some notable value for safety evaluation, we believe that for those medicinals with higher case numbers showing an unremarkable course of pregnancy, the data suggest that a preliminary positive safety statement is warranted. Regarding teratogenicity, there are insufficient data for evidence of safety. Therefore, for most Chinese medicinals, the situation remains similar to that of the majority of chemical drugs, namely, that neither the safety nor the risk during pregnancy can be definitely ascertained. Further studies on the risks of Chinese medicinals during pregnancy, chiefly epidemiological studies and documented administrations, are urgently needed to investigate the issue of possible teratogenicity.

Conflict of interest statement

The authors indicate that they do not have conflicts of interest in this work.

Acknowledgements

The authors thank Prof. Dr. Benno Brinkhaus, Institute for Social Medicine, Epidemiology and Health Economics at the Charité University Medical Center, Berlin, for valuable instructions for preparation of the manuscript, PD Dr. Christof Schaefer, Institute for Pharmacovigilance and Teratology, at the Charité University Medical Center, Berlin for reviewing the manuscript, and Dr. Velia Wortman, Nuremberg, for help with the translation of the text. The study was financially supported by the Scientific Chapter of the *Deutsche Ärztegesellschaft für Akupunktur*, Munich, Germany, the *Schweizerische Berufsorganisation für TCM*, Degersheim, Switzerland, and the *Societas Medicinae Sinensis*, Munich, Germany.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ctim.2014.08.005.

*Included studies from AR are cited in Supplementary Table 1; included studies from CR are cited in Supplementary Table 2.

References

1. Liu J. TCM exports set to rise at a healthy clip. *China Daily* 2012. ⟨http://www.chinadaily.com.cn/bizchina/2012-02/10/content_14578000.htm⟩ (date accessed 2014/07/30).
2. Hiru Corporation. *Hiru Corporation and traditional Chinese medicine market research report*. Hiru Corporation; 2010. ⟨http://www.docstoc.com/docs/45250839/Hiru-Corporation-and-Traditional-Chinese-Medicine-Market-Research⟩ (date accessed 2012/12/27).

3. WHO Commission on Intellectual Property Innovation and Public Health. *Traditional Chinese medicine could make ''health for one'' true*. WHO Commission on Intellectual Property Innovation and Public Health; 2005. ⟨http://www.who.int/intellectual property/studies/Jia.pdf⟩ (date accessed 01/12/2013).
4. Bensky D, Clavey S, Stöger E. *Chinese herbal medicine. Materia medica*. 3rd ed. Seattle: Eastland Press; 2004.
5. Hempen CH, Fischer T. *A materia medica for Chinese medicine: plants, minerals and animal products*. 1st ed. Edinburgh: Churchill Livingstone; 2009.
6. Chen JK, Chen TT. *Chinese medical herbology and pharmacology*. 1st ed. Art of Medicine Press; 2004.
7. Wang CC, Li L, Tang LY, Leung PC. Safety evaluation of commonly used Chinese herbal medicines during pregnancy in mice. *Hum Reprod* 2012;**27**:2448—56.
8. Li L, Tang LY, Man GC, et al. Potential reproductive toxicity of Largehead Atractylodes Rhizome, the most commonly used Chinese medicine for threatened miscarriage. *Hum Reprod* 2011;**26**:3280—8.
9. Li L, Dou LX, Neilson JP, Leung PC, Wang CC. Adverse outcomes of Chinese medicines used for threatened miscarriage: a systematic review and meta-analysis. *Hum Reprod Update* 2012;**18**:504—24.
10. Li L, Dou L, Leung PC, Wang CC. Chinese herbal medicines for threatened miscarriage. *Cochrane Database Syst Rev* 2012;**5**:CD008510.
11. Mitchell AA. Studies of drug-induced birth defects. In: Strom BL, editor. *Pharmacoepidemiology*. Chichester: John Wiley & Sons; 2005. p. p. 501—514.
12. Schaefer C, Spielmann H, Vetter K, Weber-Schöndorfer C. *Arzneimittel in Schwangerschaft und Stillzeit*. 8th ed. München: Urban&Fischer; 2012.
13. ICH. *ICH harmonised tripartite guideline: detection of toxicity to reproduction for medical products & toxicity to male fertility S5(R2)*; 2005. ⟨http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S5_R2/Step4/S5_R2__Guideline.pdf⟩ (last update 2005).
14. Bailey J, Knight A, Balcombe J. The future of teratology research is *in vitro*. *Biogenic Amines* 2005;**19**:97—145.
15. Chinese Pharmacopoeia Commission. *Pharmacopoeia of the People's Republic of China*, Vol. I. Beijing: China Medical Science Press; 2010.
16. Liu MS, Li XH, Zhang H, et al. [Effects of herba taxilli on the embryotic development in SD pregnant rats] (Chinese). *Zhongguo Yousheng Yu Yichuan Zazhi* 2013;**21**:105—7, 58.
17. Xu M, Tian XY, Chong LL, et al. Developmental toxicity evaluation of Rhizoma Atractylodis Macrocephalae: in-vivo and in-vitro studies (Abstract S2-20). In: Chinese Society of Toxicology, editor. *The abstracts book of the 5th National Congress of the Chinese Society of Toxicology*. Beijing: Chinese Society of Toxicology; 2009. p. 33—4.
18. Yang GY, Luo H, Liao X, Liu JP. Chinese herbal medicine for the treatment of recurrent miscarriage: a systematic review of randomized clinical trials. *BMC Complement Altern Med* 2013;**13**:320.
19. Polifka JE, Friedman JM. Clinical teratology: identifying teratogenic risks in humans. *Clin Genet* 1999;**56**:409—20.
20. Caffeine: FDA. deletion of GRAS status, proposed declaration that no prior sanction exists, and use on an interim basis pending additional study. *Fed Regist* 1980;**45**:69817—38.
21. Friedman M, Rayburn JR, Bantle JA. Developmental toxicology of potato alkaloids in the frog embryo teratogenesis assay—Xenopus (FETAX). *Food Chem Toxicol* 1991;**29**:537—47.
22. Kline BE, Von Elbe H, Dahle NA, Kupchan SM. Toxic effects of potato sprouts and of solanine fed to pregnant rats. *Proc Soc Exp Biol Med* 1961;**107**:807—9.
23. Gaffield W, Keeler RF. Induction of terata in hamsters by solanidane alkaloids derived from Solanum tuberosum. *Chem Res Toxicol* 1996;**9**:426—33.
24. Chuang CH, Doyle P, Wang JD, et al. Herbal medicines used during the first trimester and major congenital malformations: an analysis of data from a pregnancy cohort study. *Drug Saf* 2006;**29**:537—48.
25. Chuang CH, Doyle P, Wang JD, et al. Herbal medicines during pregnancy and childhood cancers: an analysis of data from a pregnancy cohort study. *Pharmacoepidemiol Drug Saf* 2009;**18**:1119—20.
26. Johns T, Sibeko L. Pregnancy outcomes in women using herbal therapies. *Birth Defects Res B Dev Reprod Toxicol* 2003;**68**:501—4.
27. Jurgens TM. Potential toxicities of herbal therapies in the developing fetus. *Birth Defects Res B Dev Reprod Toxicol* 2003;**68**:496—8.
28. Low Dog T. The use of botanicals during pregnancy and lactation. *Altern Ther Health Med* 2009;**15**:54—8.
29. Li DZ. Use of traditional Chinese herbal medicines during early pregnancy in mainland China. *Pharmacoepidemiol Drug Saf* 2007;**16**:942—3.
30. Marcus DM, Snodgrass WR. Do no harm: avoidance of herbal medicines during pregnancy. *Obstet Gynecol* 2005;**105**:1119—22.
31. EFSA Scientific Committee. Guidance on safety assessment of botanicals and botanical preparations intended for use as ingredients in food supplements. *EFSA J* 2009;**7**:1249—68.
32. Adam MP, Polifka JE, Friedman JM. Evolving knowledge of the teratogenicity of medications in human pregnancy. *Am J Med Genet C Semin Med Genet* 2011;**157**:175—82.
33. Lammer EJ. Embryopathy in infant conceived one year after termination of maternal etretinate. *Lancet* 1988;**2**:1080—1.
34. Arzneitelegramm. Schwangerschaft. Wartezeit nach Absetzen des Psoriasis-Retinoids Acitretin (NEOTIGASON). *Arzneitelegramm* 1997;**28**:58.