



รายงานฉบับสมบูรณ์

ประสิทธิภาพและความปลอดภัยของสมุนไพรที่ได้รับการส่งเสริมหลัก 5 รายการ
กวาวเครือขาว บัวบก กระจับปี่ ไพล และ ลูกประคบ การทบทวนวรรณกรรม
อย่างเป็นระบบและการวิเคราะห์ห่อภิมาณ

คณะผู้วิจัย

ดร. ปิยะเมธ ดิลกธรสกุล

รองศาสตราจารย์ ดร. ณรร ชัยญาคุณาภักษ์

รองศาสตราจารย์ ดร. ชื่นจิตร กองแก้ว

ผู้ช่วยศาสตราจารย์ ดร. ธีรพล ทิพย์พยอม

ผู้ช่วยศาสตราจารย์ ดร. สุรศักดิ์ เส้าแก้ว

ดร. ภาณุพงศ์ พุทธรักษ์

ดร. รสรินทร์ สรวมศิริ

คณะเภสัชศาสตร์ มหาวิทยาลัยนเรศวร

ได้รับทุนสนับสนุนการศึกษาวิจัยจากกองทุนภูมิปัญญาการแพทย์แผนไทย

17 กรกฎาคม 2560

บทสรุปผู้บริหาร

ปัจจุบันประเทศไทยมีค่าใช้จ่ายด้านสุขภาพสูงขึ้นในทุกๆ ปี ในปี แนวทางหนึ่งที่จะสามารถลดรายจ่ายด้านยาของประเทศลงได้คือการสนับสนุนการใช้สมุนไพรหรือการแพทย์ทางเลือกซึ่งมีต้นทุนต่ำและอาจช่วยเสริมการรักษาด้วยยาแผนปัจจุบัน

กรมพัฒนาการแพทย์แผนไทยและการแพทย์ทางเลือกได้เสนอสมุนไพรแซมเปียน 5 รายการได้แก่ กวาวเครือขาว บัวบก กระจับปี่ ไพลและลูกประคบ ที่จะทำการส่งเสริมและพัฒนาเพื่อให้สามารถผลิตเป็นผลิตภัณฑ์ออกสู่ตลาดได้ อันจะสามารถช่วยส่งเสริมให้มีการใช้ในการรักษาหรือบรรเทาอาการของโรค และอาจช่วยให้สามารถลดรายจ่ายด้านสุขภาพของประเทศลงได้

อย่างไรก็ตาม ข้อมูลด้านประสิทธิภาพและความปลอดภัยของสมุนไพรทั้ง 5 รายการนั้นยังไม่เป็นที่แน่ชัด ยังไม่มีการรวบรวมข้อมูลและสรุปอย่างเป็นระบบ โดยเฉพาะข้อมูลการศึกษาในมนุษย์ดังนั้นผู้วิจัยจึงได้จัดทำการศึกษาขึ้นเพื่อ

1. สืบค้นและรวบรวมวรรณกรรมที่ศึกษาประสิทธิภาพและความปลอดภัยของสมุนไพรที่มีโอกาสต่อยอดเป็นผลิตภัณฑ์ออกจำหน่าย 5 รายการได้แก่ กวาวเครือขาว บัวบก กระจับปี่ ไพลและลูกประคบ
2. เพื่อวิเคราะห์ห่อภิมาณถึงผลเกี่ยวกับฤทธิ์ทางประสิทธิภาพและความปลอดภัย ของสมุนไพรทั้ง 5 รายการ

จากการทบทวนวรรณกรรมอย่างเป็นระบบและการวิเคราะห์ห่อภิมาณแบ่งตามชนิดของสมุนไพรพบว่า

1. กวาวเครือขาว อาจสามารถ ลดอาการหลังหมดประจำเดือนได้ในหญิงวัยหมดประจำเดือนและสามารถเพิ่มคุณภาพของช่องคลอดได้ และอาจจะสามารถเพิ่มกระดูกได้ แต่ไม่พบข้อมูลที่เกี่ยวข้องกับการเพิ่มฮอร์โมนเพศและความปลอดภัยในการใช้
2. ยังไม่มีข้อมูลเพียงพอสำหรับการศึกษาประสิทธิภาพของ บัวบก ต่อการเพิ่มความรับรู้และความจำ แต่พบว่า อาจส่งผลต่อความสามารถในการจัดการอารมณ์โกรธและความตื่นตัว
3. กระจับปี่ อาจสามารถเพิ่มสมรรถภาพของร่างกายในแง่ของการเพิ่มพลังกำลัง การเพิ่มสมรรถภาพทางเพศของผู้ชาย และลดอาการปวดบวมได้ แต่อย่างไรก็ตาม ยังไม่พบข้อมูลสนับสนุนการใช้กระจับปี่ต่อ
4. ไพลครีม ที่ความเข้มข้น 14% อาจสามารถลดอาการปวดกล้ามเนื้อ และอาการอักเสบของข้อได้
5. ลูกประคบไทย อาจ มีประสิทธิภาพในการลดอาการปวดท้องจากการคลอด (labor pain) และลดระยะเวลาในการหลั่งน้ำนมในคุณแม่หลังคลอดได้ แต่ไม่พบข้อมูลยืนยันด้านประสิทธิภาพในการลดอาการปวดเข่าและอาการปวดกล้ามเนื้อที่แตกต่างจากการประคบร้อนและการใช้ยาในกลุ่ม non-steroidal anti-inflammatory drugs

อย่างไรก็ตาม การทบทวนวรรณกรรมอย่างเป็นระบบและการวิเคราะห์ห่อภิมาณของสมุนไพรทั้ง 5 ชนิดนี้ ยังคงมีข้อจำกัดที่สำคัญ เช่น การศึกษาส่วนใหญ่ที่ได้รับการคัดเข้ายังมีคุณภาพต่ำ มีขนาดของผู้เข้าการศึกษาน้อย หรือวัดผลได้ไม่เหมาะสม เป็นต้น ดังนั้นการนำการศึกษาไปใช้ควรเป็นไปอย่างระมัดระวัง และคำนึงถึงผลกระทบของการนำผลการศึกษานี้ไปใช้เป็นสำคัญ

กาวาเครือขาว

การสืบค้นข้อมูลอย่างเป็นระบบเพื่อศึกษาประสิทธิผลและความปลอดภัยของการใช้กาวาเครือขาวในผู้หญิงวัยหมดประจำเดือน โดยทำการสืบค้นข้อมูลจากฐานข้อมูลทั้งหมด 12 ฐานข้อมูล ได้แก่ AMED, CINAHL, Cochrane Central Register of clinical trials, EMBASE, Health Science Journals in Thailand, PubMed, Thai Index Medicus, Thai Library Integrated System, Thai Medical Index, Thai Thesis Database, WHO registry, และ www.clinicaltrial.gov จนถึงเดือนตุลาคม 2016 พบการศึกษาทั้งหมด 7 การศึกษาจากผู้เข้าร่วมการศึกษารวม 233 คน โดยแบ่งเป็นการศึกษาแบบทดลองและมีกลุ่มควบคุม 3 การศึกษา (Randomized control clinical trial; RCT) และการศึกษาที่ไม่มีกลุ่มควบคุม 4 การศึกษา

การวิเคราะห์ประสิทธิผลของกาวาเครือขาวแบ่งออกเป็น 3 ผลลัพธ์ ได้แก่ 1) ประสิทธิภาพของกาวาเครือขาวต่ออาการหลังหมดประจำเดือน 2) ประสิทธิภาพของกาวาเครือขาวต่อฮอร์โมนเพศ และ 3) ประสิทธิภาพของกาวาเครือขาวต่อกระดูก

ประสิทธิผลของกาวาเครือขาวต่ออาการหลังหมดประจำเดือน

การศึกษาประสิทธิผลของกาวาเครือขาวต่ออาการหลังหมดประจำเดือนโดยวัดจากเครื่องมือ itemized modified Greene climacteric scale (MGCS) พบว่า กาวาเครือขาวขนาด 25 – 100 มิลลิกรัมของกาวาเครือขาวต่อวัน สามารถลดคะแนน MGCS ได้ประมาณร้อยละ 50 และสามารถลดอาการร้อนวูบวาบ (hot flushes) ได้ประมาณร้อยละ 78 เมื่อเทียบกับก่อนการได้กาวาเครือขาว โดยกาวาเครือขาวในขนาด 25 และ 50 มิลลิกรัมสามารถลดอาการหลังหมดประจำเดือนได้ไม่แตกต่างกัน นอกจากนี้ยังพบว่า กาวาเครือขาวสามารถลดอาการหงุดหงิดง่ายและการนอนหลับผิด ได้เช่นกัน

การศึกษาประสิทธิผลของกาวาเครือขาวต่อสุขภาพของช่องคลอด พบว่า สุขภาพของช่องคลอดวัดโดย vaginal health-index, maturation index, maturation value เพิ่มขึ้นในสัปดาห์ที่ 12 และ 24 หลักจากได้รับประทานกาวาเครือขาวเปรียบเทียบกับก่อนได้รับประทานกาวาเครือขาวและเปรียบเทียบกับยาหลอก

ประสิทธิผลของกาวาเครือขาวต่อฮอร์โมนเพศ

การศึกษาประสิทธิผลของกาวาเครือขาวต่อฮอร์โมนเพศ พบว่า ยังไม่มีข้อมูลที่ชัดเจนของประสิทธิผลของกาวาเครือขาว, estradiol, luteinizing hormone, และ follicle stimulating hormone ในกระแสเลือด

ประสิทธิผลของกาวาเครือขาวต่อกระดูก

การศึกษาประสิทธิผลของกาวาเครือขาวต่อกระดูก พบว่า กาวาเครือขาวมีสามารถลดการทำงานของ osteoclast การดึงเกลือแร่ออกจากกระดูก (demineralization) และยับยั้งการทำงานของ bone-specific alkaline phosphatase ซึ่งอาจส่งผลให้สามารถเพิ่มมวลกระดูกได้

ความปลอดภัยของกาวาเครือขาว

อาการไม่พึงประสงค์ของกาวาเครือขาวที่พบได้บ่อย ได้แก่ อาการเจ็บคัดเต้านมและอาการไม่พึงประสงค์ทางช่องคลอด โดยที่อัตราการเกิดอาการไม่พึงประสงค์ดังกล่าวไม่แตกต่างจากยาหลอกและ conjugated equine estrogen

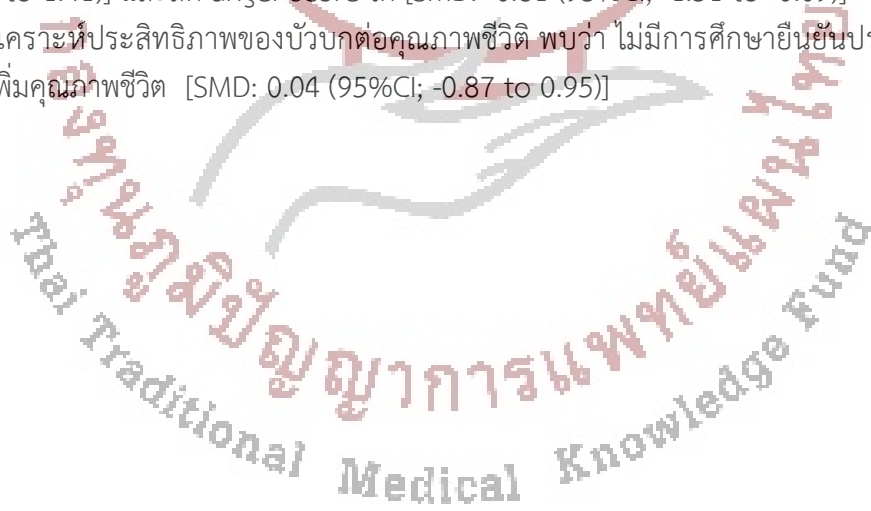
บัวบก

การสืบค้นข้อมูลอย่างเป็นระบบเพื่อศึกษาประสิทธิภาพและความปลอดภัยของการใช้บัวบกต่อการรับรู้และความจำ (cognitive function) อารมณ์ (mood) และคุณภาพชีวิต โดยทำการสืบค้นข้อมูลจากฐานข้อมูลทั้งหมด 13 ฐานข้อมูล ได้แก่ AMED, CINAHL, Cochrane Central Register of clinical trial, EMBASE, PubMed, Psyninfo, Science direct, Scopus, www.clinicaltrial.gov, ThaiLis, Thai Index Medicus, Thai Medical Index, and Thai Thesis Database จนถึงเดือนกันยายน 2016 พบการศึกษาทั้งหมด 11 การศึกษาแบบทดลองและมีกลุ่มควบคุม (Randomized control clinical trial; RCT) โดย 5 การศึกษาเปรียบเทียบระหว่างการใช้บัวบกเดี่ยวกับยาหลอก และอีก 6 การศึกษาเปรียบเทียบการใช้บัวบกร่วมกับสมุนไพรอื่นๆ กับสมุนไพรชนิดอื่นๆ

การวิเคราะห์ประสิทธิภาพของบัวบกต่อการเพิ่มความรับรู้และความจำ (cognitive function) พบว่ายังไม่มีการศึกษายืนยันชัดเจนในแง่ของประสิทธิภาพของการรับประทานบัวบกต่อการเพิ่มความรับรู้และความจำทั้งในด้าน overall cognitive function, attention and concentration, executive function, information processing speed, language, visuospatial skill, working memory, verbal memory, visual memory มีเพียงการวิเคราะห์เดียวที่แสดงผลของการรับประทานบัวบกร่วมกับสมุนไพรอื่นๆ ที่มีผลเพิ่ม executive function ในแง่ของการลดเวลาในการแก้ปัญหา (time of problem solving) [standardized mean difference (SMD): -0.77 (95% confidence interval (CI); -1.26 to -0.27]

การวิเคราะห์ประสิทธิภาพของบัวบกต่ออารมณ์ พบว่า บัวบกสามารถเพิ่ม alert score [SMD: 0.71 (95%CI; 0.01 to 1.41)] และลด anger score ได้ [SMD: -0.81 (95%CI; -1.51 to -0.09)]

การวิเคราะห์ประสิทธิภาพของบัวบกต่อคุณภาพชีวิต พบว่า ไม่มีการศึกษายืนยันประสิทธิภาพของบัวบกต่อการเพิ่มคุณภาพชีวิต [SMD: 0.04 (95%CI; -0.87 to 0.95)]



กระชายดำ

การสืบค้นข้อมูลอย่างเป็นระบบเพื่อศึกษาประสิทธิภาพทางคลินิกและความปลอดภัยของการใช้กระชายดำในทุกๆข้อบ่งชี้ โดยทำการสืบค้นจากฐานข้อมูลทั้งหมด 12 ฐานข้อมูล ได้แก่ PubMed, EMBASE, Cochrane Central Register of Clinical Trials, CINAHL, AMED, WHO Registry, www.clinicaltrial.gov, Health Science Journals in Thailand, Thai Library Integrated System, Thai Thesis Database, Thai Index Medicus, and Thai Medical Index จนถึงเดือนมกราคม 2016 พบการศึกษาทั้งหมด 7 การศึกษาแบบทดลองและมีกลุ่มควบคุม (Randomized control clinical trial; RCT) ที่เข้าเกณฑ์การคัดเลือก โดย 6 การศึกษาจากทั้งหมด 7 การศึกษาเป็นการศึกษาในประเทศไทย

ข้อบ่งชี้ที่ได้รับการศึกษาจากการศึกษาที่ได้รับการคัดเลือกทั้งหมด 4 ข้อบ่งชี้ ได้แก่ 1) ประสิทธิภาพต่อการเพิ่มสมรรถภาพของร่างกาย (physical/exercise performance) 2) ประสิทธิภาพต่อการเพิ่มการตอบสนองการแข็งตัวของอวัยวะเพศชาย (erectile response) 3) ประสิทธิภาพต่อการลดปวด (pain indicator) และ 4) ประสิทธิภาพต่อการใช้กำลัง (energy expenditure)

ประสิทธิภาพของกระชายดำต่อการเพิ่มสมรรถภาพของร่างกาย

กระชายดำสามารถเพิ่มสมรรถภาพของร่างกายได้ ในแง่ของความแรงของการจับมือ (hand grip strength) ทั้งมือขวาและมือซ้าย [standardized mean difference (SMD); - .44 (95%CI; 0.02 – 0.86) สำหรับมือขวา, SMD; 0.49 (95%CI; 0.07 – 0.91) สำหรับมือซ้าย]

ประสิทธิภาพของกระชายดำต่อการเพิ่มการตอบสนองการแข็งตัวของอวัยวะเพศชาย

กระชายดำที่ขนาด 90 มิลลิกรัมต่อวันสามารถลดระยะเวลาในการตอบสนองต่อสิ่งเร้าสำหรับการแข็งตัวของอวัยวะเพศชายได้ และยังสามารถเพิ่มความยาวและความใหญ่ของอวัยวะเพศได้ทั้งในภาวะพัก (resting state) และในภาวะแข็งตัว (erection state) เมื่อเปรียบเทียบกับยาหลอก แต่ไม่พบความเปลี่ยนแปลงของระดับฮอร์โมนในร่างกาย อย่างไรก็ตาม การศึกษาในประสิทธิภาพนี้ พบเพียง 1 การศึกษาที่มีกลุ่มตัวอย่างเพียง 45 คนเท่านั้น การแปลผลการศึกษาอาจต้องทำอย่างระมัดระวัง

ประสิทธิภาพของกระชายดำต่อการลดปวด

กระชายดำสามารถลดความรุนแรงในการปวดข้อเข่าได้เมื่อใช้รักษาเป็นระยะเวลา 4 สัปดาห์ รวมทั้งสามารถลดการบวมของข้อเข่า (circumference of knee joint) และเพิ่มช่วงการเคลื่อนไหวของข้อเข่า (range of motion of knee joint) ได้เมื่อเทียบกับก่อนการรับประทานกระชายดำ อย่างไรก็ตาม การศึกษาในประสิทธิภาพนี้ พบเพียง 1 การศึกษาเท่านั้น การแปลผลการศึกษาอาจต้องทำอย่างระมัดระวัง

ประสิทธิภาพของกระชายดำต่อการใช้กำลัง

กระชายดำในขนาด 100 และ 180 มิลลิกรัมต่อวันสามารถเพิ่มอัตราการใช้กำลังได้ที่เวลา 30 และ 60 นาที เมื่อเปรียบเทียบกับก่อนการรับประทานกระชายดำ แต่เมื่อเปรียบเทียบกับยาหลอก ไม่พบความแตกต่างของอัตราการใช้กำลังที่เพิ่มขึ้น อย่างไรก็ตาม การศึกษาในประสิทธิภาพนี้ พบเพียง 1 การศึกษาเท่านั้น การแปลผลการศึกษาอาจต้องทำอย่างระมัดระวัง

ไพล

การสืบค้นข้อมูลอย่างเป็นระบบเพื่อศึกษาประสิทธิภาพทางคลินิกของการไพลในทุกๆข้อบ่งใช้ โดยทำการสืบค้นจากฐานข้อมูลทั้งหมด 12 ฐานข้อมูล ได้แก่ AMED, CENTRAL, CINAHL, ClinicalTrial.gov, EMBASE, Health Science Journals, MEDLINE, Thai Index Medicus, Thai Medical Index, Thai Thesis Database, Thai LIS and WHO registry จนถึงเดือน กุมภาพันธ์ 2017 พบการศึกษาทั้งหมด 6 การศึกษา โดยมี 4 การศึกษาเป็นการศึกษาแบบทดลองและมีกลุ่มควบคุม (Randomized control clinical trial; RCT) และ 2 การศึกษาเป็นการศึกษาแบบ Quasi-experimental

ข้อบ่งใช้ที่ได้รับการศึกษาจากการศึกษาที่ได้รับการคัดเข้าทั้งหมด 3 ข้อบ่งใช้ ได้แก่ 1) ประสิทธิภาพของไพลต่อการลดอาการปวด (pain reduction) 2) ประสิทธิภาพของไพลต่อการลดสิว และ 3) ประสิทธิภาพของไพลต่อการต้านการอักเสบ

ประสิทธิภาพของไพลต่อการลดอาการปวด

การศึกษาประสิทธิภาพของไพลต่อการลดอาการปวดกล้ามเนื้อ พบว่า การศึกษาส่วนใหญ่ บ่งชี้ไปในทางเดียวกันว่า ครีมไพลที่ความเข้มข้น 14% สามารถลดอาการปวดกล้ามเนื้อได้เมื่อเปรียบเทียบกับก่อนการใช้ครีมไพลและยาหลอก แต่ไม่พบประสิทธิภาพของครีมไพลที่ความเข้มข้น 7% ต่อการลดอาการปวดกล้ามเนื้อ

การศึกษาประสิทธิภาพของไพลต่อการปวดจากข้อเท้าแพลง พบว่า ครีมไพลที่ความเข้มข้น 14% สามารถลดอาการปวดจากข้อเท้าแพลงได้เมื่อเทียบกับยาหลอกที่เวลา 5 – 6 วัน

การศึกษาประสิทธิภาพของไพลต่อการปวดคอและไหล่ พบว่า การใช้การนวดประคบด้วยไพลสามารถลดอาการปวดคอและไหล่ได้เมื่อเทียบกับก่อนการนวดประคบ แต่ไม่แตกต่างกับการนวดประคบแบบไทย (Thai traditional massage)

ประสิทธิภาพของไพลต่อการลดสิว

การศึกษาประสิทธิภาพของไพลเจลที่ความเข้มข้น 1% ต่อการรักษาสิวนิดไม่รุนแรง (mild to moderate acne) พบว่า ไพลเจล 1% สามารถรักษาสิวได้ไม่แตกต่างจากยาหลอก ทั้งในสิวนักเสบและสิวนักเสบ

ประสิทธิภาพของไพลต่อการต้านการอักเสบ

การศึกษาประสิทธิภาพของ ยาเม็ดไพลแห้งขนาด 500 มิลลิกรัมเปรียบเทียบกับยาคลอร์เฟนิรามีนในการลดการอักเสบ ซึ่งวัดจากขนาดของผื่นก่อนและหลังการใช้ยา พบว่า ยาเม็ดไพลแห้งขนาด 500 มิลลิกรัมลดอาการอักเสบได้น้อยกว่ายาคลอร์เฟนิรามีน

ลูกประคบไทย

การสืบค้นข้อมูลอย่างเป็นระบบเพื่อศึกษาประสิทธิผลทางคลินิกของการใช้ลูกประคบในทุกๆข้อบ่งใช้ โดยทำการสืบค้นจากฐานข้อมูลทั้งหมด 12 ฐานข้อมูล ได้แก่ AMED, CINAHL, Cochrane Central Register of clinical trial, EMBASE, Health Science Journals in Thailand, PubMed, Thai Index Medicus, Thai Library Integrated System, Thai Medical Index, Thai Thesis Database, WHO registry, and www.clinicaltrial.gov พบการศึกษาทั้งหมด 13 การศึกษาจากผู้เข้าร่วมการศึกษารวม 778 คน โดยมี 1 การศึกษาเป็นการศึกษาแบบทดลองและมีกลุ่มควบคุม (Randomized control clinical trial; RCT) และ 12 การศึกษาเป็นการศึกษาแบบ Quasi-experimental

ข้อบ่งใช้ที่ได้รับการศึกษาจากการศึกษาที่ได้รับการคัดเข้ามีทั้งหมด 4 ข้อบ่งใช้ ได้แก่ 1) ประสิทธิภาพของลูกประคบไทยต่อการปวดข้อเข่า (knee osteoarthritis pain) 2) ประสิทธิภาพของลูกประคบไทยต่อการปวดกล้ามเนื้อ (muscle pain) 3) ประสิทธิภาพของลูกประคบไทยต่อการปวดท้องคลอด (labor pain) และ 4) ประสิทธิภาพของลูกประคบไทยต่อการหลั่งน้ำนม (lactation)

ประสิทธิผลของลูกประคบไทยต่อการปวดข้อเข่า

ลูกประคบไทยไม่สามารถอาการปวดเข่าได้เมื่อทำการวัดเป็นคะแนนการปวด ทั้งเมื่อเปรียบเทียบกับยา non-steroidal anti-inflammatory drugs (NSAIDs) [SMD: 0.351 (95%CI; -0.091 to 0.793)], เปรียบเทียบกับการออกกำลังกายเข่า (knee exercise) [mean difference (MD): 0.660 (95%CI; -0.839 to 2.160)], เปรียบเทียบกับการประคบร้อน (hot compress) [SMD: 0.254 (95%CI; -0.161 to 0.669)] และเมื่อเปรียบเทียบกับตัวเปรียบเทียบใดก็ได้ (combined comparator) [SMD: 0.419 (95%CI; -0.004 to 0.842)]

นอกจากนี้เมื่อทำการวิเคราะห์ความสามารถในการทำกิจกรรมสำหรับผู้ป่วยข้อเข่าเสื่อม พบว่า ลูกประคบไทยไม่เพิ่มความสามารถในการทำกิจกรรมเมื่อเปรียบเทียบกับ NSAIDs, การออกกำลังกายเข่า, การประคบร้อน หรือตัวเปรียบเทียบใดก็ได้

ประสิทธิผลของลูกประคบไทยต่อการปวดกล้ามเนื้อ

ลูกประคบไทยไม่สามารถลดอาการปวดกล้ามเนื้อได้เมื่อเปรียบเทียบกับ (NSAIDs) [MD: 0.300 (95%CI; -0.539 to 1.139)], การประคบร้อน (hot compress) [MD: 0.525 (95%CI; -0.842 to 1.839)] และตัวเปรียบเทียบใดก็ได้ (combined comparator) [MD: 0.4135 (95%CI; -0.574 to 0.845)]

ประสิทธิผลของลูกประคบไทยต่อการปวดท้องคลอด

ลูกประคบไทยสามารถลดอาการปวดหลัง หลังจากการคลอด ได้เมื่อเปรียบเทียบกับการรักษามาตรฐาน (usual care) และสามารถลดอาการปวดท้องคลอด (labor pain) ในแม่ที่ตั้งครรภ์ครั้งแรกได้

ประสิทธิผลของลูกประคบไทยต่อการหลั่งน้ำนม

ลูกประคบไทยสามารถลดระยะเวลาในการหลั่งน้ำนมหลังจากคลอดได้เมื่อเปรียบเทียบกับการรักษาปกติ (routine program) [MD: -394.425 นาที (95%CI; -620.084 to -168.766)]

เรื่องที่ 1: Efficacy and safety of *Pueraria mirifica* for menopause: A systematic review of clinical trials and the way forward

Introduction

For many middle-aged women, ovarian degeneration causes intolerable menopausal symptoms including vasomotor dysfunction (hot flashes, sweating), sleep disorders (insomnia, waking and erratic sleep), depression, sexual dysfunction (reduced libido, urogenital disorders), and musculoskeletal pain. Vasomotor symptoms affects 57% of women, begins before menopause and continues for 1 to >15y (mean 4.5y), but is shorter in Asians, and more prolonged in African Americans, the obese, premature menopausal, and the stressed (1). Estrogen replacement is the rational treatment but the 2002 Women Health Initiative trial confirmed several adverse actions as major risks, especially thrombosis and neoplasms. Prescription rates plummeted leaving women with no clear alternative so many turned to phytoestrogens. But, the common phytoestrogens, isoflavones and lignans, have unproven effectiveness against menopause symptoms (2). This perhaps accords with their predominant action on estrogen receptor (ER)(3).

Legumes commonly express isoflavones and also the more potent estrogens, coumestrols. *Leguminosae Pueraria* genus comprises 15-20 species that are invasive woody climbing perennial beans indigenous to south-east Asia but genetically diverse(4). They are used for fodder (foliage), human food and medicines (tubers)(5). Apart from *P. montana* (Kudzu), most *pueraria* species occupy niches and one of these is *Pueraria candollei*, is localized to Bangladesh, Assam, Myanmar, and Thailand. Several varieties and cultivars been studied for their estrogenicity, especially var. *mirifica* (Airy Shaw & Suvat.) Niyomdham (commonly termed *P. mirifica*, or in the Thai language, white kwao kuer (various spellings), sometimes Thai kudzu. It is restricted to a few deciduous hilly jungles in north and western Thailand (6, 7) but has been over-exploited and become endangered. Accordingly, it is extensively farmed for medicines. *P. candollei* var *candollei* also appears to contain estrogens, while several cultivars have been tested in preclinical studies.

Traditional *P. mirifica* uses were largely restricted to regions of Thailand where the wild plant was gathered. Its action emphasised rejuvenation and good health in older women (and men), including increased energy, infection control and parasite tolerance, healthy skin and hair growth, cognitive function, reduced cataracts, less breast sag, improved sleep, blood 'tonic', but contra-indicated for younger adults. Adverse effects were muscle/joint pain but alleviated by cold showers (8). There seems to be little emphasis on treating the psychiatric/vasomotor postmenopausal symptoms, probably reflecting their lower incidence in ethnic Asians and healthier diets of ancients in these Thai regions (8-10). So, these indigenants appear to have recognized the pleiotropic actions in old age without knowledge of hormones. These changes

accord with currently appreciated postmenopausal tissue changes including increased cardiovascular risks, bone degeneration, cognitive decline, metabolic disease, immune attenuation, and deterioration of skin and hair.

P. mirifica, especially its tubers contain glycosides of genistein, and daidzein, including the rare puerarin and mirificin, the unique kwakhurin and mirificoumastans (5), and notably miroesterol and its congeners. These are potent estrogen receptor (ER) agonists in vitro and in animals (11, 12). Pre-clinical studies suggest miroesterols are responsible for major clinical actions (13) although there are no supporting pharmacokinetic or metabolic studies. Their ER actions would be expected to replicate standard estrogen replacement therapy in alleviating post-menopausal symptoms.

However, early studies standardized *P. mirifica* preparations by HPLC of isoflavones whose total contents varied 50-fold between season, cultivar, and region(14). Similar wide variations were found using bioassay (rat vaginal cornification) in *P. mirifica* sampled by province(15). Miroestrol in *P. mirifica* also varied by region and variety (Kitisripanya et al., 2017). In commercial *P. mirifica* products, oxymiroestrol concentrations varied ~70-fold(16) while some commercial formulations claiming to be *P. mirifica* contained no *P. mirifica* (17). Genomic studies have detected sequence differences in batches of *P. mirifica* (18, 19) which may contribute to content inconsistencies. Thus, chromatographic finger-prints of relevant active constituents are mandatory for replicatability.

Currently, *P. mirifica* products are globally promoted for breast enlargement, cosmetics, and menopausal symptoms (Google search of 'pueraria' yields results dominated by PM). Purveyors of *P. mirifica* cite relevant clinical trials in glowing terms to support its efficacy while many products carry appropriate warnings, others describe it as "completely safe" again citing the scientific literature. Furthermore, internet blogs on *P. mirifica* reveal many consumers using 10x recommended doses in an attempt to titrate an effective dose regime. Clearly, users need an independent and unbiased expert assessment of *P. mirifica* through a systematic review enabling them, or their medical practitioner, to make an informed risk/benefit analysis. Such a review appears not to have been undertaken and is therefore timely. Accordingly, we have conducted such a review to evaluate the clinical efficacy of *P. mirifica* to alleviate menopausal symptoms. This study should provide an unbiased summary of the efficacy and safety of *P. mirifica* on menopausal women as discovered by systematic review in order to create a safer and reproducible clinical profile.

Materials and Methods

This systematic review was conducted according to the Cochrane Collaboration framework guidelines (20), and reporting followed the PRISMA Statement (21).

Search strategies and study selection

The following databases were searched for original research articles from their inception to October 2016: AMED, CINAHL, Cochrane Central Register of clinical trials, EMBASE, Health Science Journals in Thailand, PubMed, Thai Index Medicus, Thai Library Integrated System, Thai Medical Index, Thai Thesis Database, WHO registry, and www.clinicaltrial.gov. Strategic search terms used scientific and common synonyms of *P.mirifica* or “kwao krua kao” (Thai for *P. mirifica*). Reference sections of all retrieved full articles were reviewed for potential studies not indexed in the above databases.

Clinical studies were included if they met the following inclusion criteria: 1) conducted on humans; 2) investigated therapeutic effects of *P. mirifica* on menopause. CK screened all the titles and abstracts to determine whether the studies assessed the therapeutic effects of *P. mirifica*. Full-text articles of the potential studies were reviewed by CK and NS and the included studies were subsequently reviewed by CK, NS, and TD. When uncertainties regarding the eligibility occurred, they were resolved by discussion.

Data extraction

Data extraction was undertaken by CK using a data extraction form and cross-checked by TD and CNS according to the CONSORT statement for reporting herbal medicinal interventions (22). The data extracted and reported included: study design; number and characteristics of participants; age of participants; herbal preparations; standardization of preparations; treatment protocols; and outcomes. Outcomes of interest related to menopause were menopausal symptoms; plasma estradiol and FSH/LH; and adverse events.

Quality assessment of included studies

The studies were assessed for methodological quality by CK, CNS, and TD using the Cochrane revised tool for risk of bias tool (RoB 2.0) for individually randomized, parallel group trials(20). This tool assesses biases due to randomization process; to deviations from intended intervention; missing outcome data; outcome measurements, selective reported of results, and overall bias. Each study was classified as having low risk (low risk of bias for all domains), high risk (high risk of bias for one or more domains), or some concerns (unclear risk of bias for one or more key domains). Disagreements between the reviewers were settled through discussion and consensus.

In addition to RoB, we used our own quality 22 quality indicators emphasising methodological validation, reliability, repeatability, ethical standards, monitoring, safety, and characterisation of the herbal medicine intervention (assessed by CK and CNS). This was based on the CONSORT checklist, its herbal medicines annex (22) and monitoring (Table 3). To keep its size manageable, some related parameters were pooled into a single domain. Measurement used a three-point score: =0 (domain parameters not performed or not reported), =0.5 (domain partly fulfilled), =1, (domain fully or almost fully adhered to). If a domain was not or largely *not applicable*, that domain was omitted in calculating the final score. The overall score gives a

general measure of quality. However, a domain of a particular study may have an extreme effect its value. Thus a participant attrition rate of 80% would suggest a serious problem with intervention, or no placebo group where placebo effects are likely to high, could nullify trial outcomes. There were also some domain parameters, e.g., trial registration, lodging raw data, which are unlikely to have been performed but included to create a generic tool for intervention trials.

All included studies were assessed by CK and CNS and disagreements resolved through discussion and consensus.

Analyses

Descriptive statistics were used in reporting the outcome data. The results are presented using a narrative approach.

Results

Search results

A total of 1060 articles were identified. Of these, 1052 were excluded because they were duplicated (227 articles), investigated *Butea superb* Roxb (red kwao krua) (68 studies), or other plant species (37 studies), not relevant to *P.mirifica* (484 articles), not clinical studies of *P.mirifica* in humans (234 articles), studied cosmetic effects of *P.mirifica* (2 articles). Two US patent applications containing identical data were excluded because participants were non-menopausal or included in Muangman and Cherdshewasart (23). Eight studies met our inclusion criteria (3, 23-29) (Figure 1, Table 1). Manonai, Chittacharoen (27), Manonai, Chittacharoen (28) were the same study but analysed different end-points. Thus there were 7 studies in total.

None of the conference abstracts found met our criteria. In 'grey' sources available in Thailand, a patent application and a Japanese study on premenopausal women but neither met our inclusion criteria.

Study characteristics

Of the seven included clinical studies, two were randomized placebo controlled trials (3, 27); one study was a head-to-head comparison with conjugated equine estrogen (CEE) with/without medroxyprogesterone acetate (MPA) but no placebo group (24, 25); four other studies also had no control groups (23-26, 29).

Primary end-points were MGCS (4 studies), urogenital tract function (1), lipids (2 studies), and undefined (1 study).

P.mirifica was dosed as capsules in four studies (23, 26-28) or tablets (3, 29), but not reported in the remaining 2 studies (24, 25). Five studies used powdered dried tubers of *P.mirifica* (3, 23, 27-29), but not reported in the others (24-26) but probably similar as the previous five. Muangman and Cherdshewasart (23) used *P. mirifica* var "Wichai III" at 200mg/day; the others used variety "Airy Shaw and Suvatabandhu" as explicitly stated or surmised from the literature

and orally administered at 20 to 100 mg/day (Table 1). The included studies were conducted on middle-aged female participants who were perimenopausal or recently postmenopausal. Intervention durations were 8 to 48 weeks.

Participant characteristics were women with self-reported menopausal symptoms seeking treatment at hospital menopause clinics or outpatient departments, except Okamura (2008) who used high serum FSH, and Manonai, Chittachoen (27) who recruited by advertisement for women having genital tract problems that were then verified by pelvic examination. High serum estradiol (~76 uM) and low FSH suggested a peri- or transitional menopausal state when estrogen levels are chaotic (Prior, 2005). Although some studies were moderately powered, others (24, 27) subdivided cohorts to determine dose-effectiveness, but the groups were too small to resolve differences. Studies used different participant characteristics, different end-points, intervention times, incomparable *P. mirifica* preparations and different protocols (positive or negative controls, no concurrent controls, or no base-line measures).

Quality of included studies

The overall risks of bias for all seven trials was generally high or of some concern (Table 2). However, in reaching the assessment in each domain involves several decision points based on generally poor reporting (see below).

Our CONSORT reporting assessment was grouped into seven domain classes (Table 3):

Authorisation: All but one study stated the Declaration of Helsinki (but not version), and the IRB giving approval but not the approval reference. All studies acknowledged the funding agency but only Virojchaiwong, Suvithayasiri (29) declared independence from funder. Manonai, Chittachoen (28) had help with chemical analyses from the funder, and three other studies imply that the funding agency had minor roles.

Designs: All studies had clear aims and appropriate end-points except one where aims were vague (23). Secondary end-points were often stated. Intervention duration was 2-12 months which was probably enough time to demonstrate effects. Only Manonai, Chittachoen (27) and Okamura, Sawada (3) used placebo groups but studies having no placebos seriously compromised definitive conclusions. Only Virojchaiwong, Suvithayasiri (29) justified cohort sizes and performed power calculations but did not account for placebo effects. For MGCS end-points without SD/SEMs, discerning under-powering was not possible. Chandeying and Lamlertkittikul (24) described their trial as open-ended making definition of duration unclear.

Participants: Inclusion criteria were vague “menopausal women were recruited” (3), menopausal women without gynecological cysts (23) while other studies had inclusion criteria appropriate for primary end-points. Except Okamura, Sawada (3), exclusion criteria were noted but the number of criteria varied between studies. For participant baseline characteristics, only Manonai, Chittachoen (27) tabulated these by intervention group while others had rudimentary information or reported nothing.

Allocation: The Muangman and Cherdshewasart (23) study was excluded here. Otherwise studies mentioning randomization were awarded “Yes/No” or NR. Differences in baseline primary end-points between groups suggested randomization was real. Participants were blinded in trials with placebo groups, but in groups comparing doses that claimed blinding, this was not real because presumably participants knew they were having treatment. For blinding of staff, reporting “double blinded” alone was inadequate. Only Manonai, Chittacharoen (27) reported this domain in detail. No study described raw data locking to prevent data manipulation.

Herbal characteristics: No study gave taxonomic or genetic information directly nor links to other publications, nor brand, etc., that could lead to characterization of the preparation. Two studies described preparation of the dried root powder. No study displayed or referenced any chromatographic ‘finger printing’. Two studies listed some ingredients, one as miroestrol and total isoflavones (3) by referring to another paper. Another study referred to a commercial laboratory with isoflavone contents but no methodology (29)

Monitoring: Only two studies attempted to assess adherence to treatment, by pill counting. All reported drop-outs/withdrawals with causes (Table 5). For safety monitoring most studies were quite thorough (lipid profiles, liver enzymes and other hepatic markers and haematology (23-26, 29). Manonai, Chittacharoen (27) also used pap smears, “endometrial surveillance”, and mammograms. Okamura, Sawada (3) reported no safety monitoring. All reported harms, but no study attempted to quantify their intensity, and there was no reports that these were followed up after treatment ceased.

Data handling: All studies losing participants accounted for lost participants but not analysed by ‘intention-to-treat’. No other data variances were reported. Data analyses were appropriate. However, the extent of analysis and excessive data pooling hindered any attempt to understand selective effects or to identify trends. Thus each menopausal woman does not present with a MGCS, but has a set of unique symptoms that often change during menopause. So, she needs to judge efficacy for her particular situation. Means without SD/SEM/CIs were common. Reporting arbitrary p-value cutoffs rather than actual p-values prevents readers from identifying trends. For example, the data for PM v. CEE or suggests a trend that CEE may have been more efficacious, especially for the important vasomotor symptoms. Any difference may be functionally insignificant, but if there are large placebo effects which are likely for MGCS (see discussion), these differences become clinically important. Manonai, Chittacharoen (27) designed three subgroups by dose but inexplicably pooled all data denying readers the opportunity make their own judgments. The raw data would have enabled readers to assess variations in individual participants but not available for any study.

Thus, the qualities of design, reporting, and analyses were generally poor in these publications.

Effects of P. mirifica

Intervention end-points of *P. mirifica* were: (A) menopausal symptoms; (B) reproductive hormones; and (C) 'other' (Table 4).

P. mirifica reduced menopause symptoms

Changes in different categories of symptoms were measured by the itemized modified Greene climacteric scale (MGCS). In the head-to-head trial between *P. mirifica* and CEE±MPA, MGCS scores in both arms were similar (25). In this and three other studies without placebos (24-26, 29), overall climacteric scores were reduced after 25-100 mg/day *P. mirifica* by ~50% compared to base-line (Table 4) but a ~78% reduction for hot flushes (the lack of detailed data reporting prevented any statistical significance being attached to this value). MGCS reductions were similar comparing 25 and 50 mg/day *P. mirifica* (24, 29). Muangman and Cherdshewasart (23) showed reduced hot flushes, frustration, and sleep disorders.

One RCT study reported *P. mirifica* actions on vaginal symptoms and morphology, as a total score of vaginal health index, or maturation value (27). Vaginal health index, evaluated by scoring vaginal parameters (moisture, fluid volume, elasticity, epithelial integrity, and pH, on scale 1 (poorest) to 5 (best) with *P. mirifica* treatment were increased at weeks 12 and 24 compared to week 0 and placebo ($p < 0.05$) (27). Vaginal maturation value and maturation index, representing vaginal cytology increased at 12 and 24 weeks of *P. mirifica* but not with placebo. Some urinary tract dysfunctions were reduced by *P. mirifica* in Lamlertkittikul and Chandeying (26), incorporated into the MGCS score (29), while other studies showed no effect on the urogenital tract (24, 25). Manonai, Chittacharoen (27) demonstrated improved genital tract function but not of the urinary tract.

Influences on reproductive hormones

In four studies, both RCT (Manonai et al., 2008) and clinical trials without placebos (Chandeying and Sangthawan, 2007) (Chandeying and Lamlertkittikul, 2007) (Virojchaiwong et al., 2011) *P. mirifica* had no clear effect on serum estradiol. In Okamura, Sawada (3) post-menopausal levels of ~70 mIU/ml for FSH was reversibly reduced by *P. mirifica*, which accords with an action mimicking estrogenic feedback on the anterior pituitary thereby inhibiting gonadotrophin release. In three studies (24-26), there were no consistent changes in serum LH or FSH with *P. mirifica*. However, baseline values for LH (~18 mIU/ml) and FSH (~38 mIU/ml) for (24) and (25) were between the normal and postmenopausal ranges and being at premenopausal levels in Lamlertkittikul and Chandeying (26) (Table 4).

Effect on bone

Manonai, Chittacharoen (27) measured the bone specific isoform suggesting an ER-mediated inhibition of osteoclast activity and reduced demineralisation. Okamura, Sawada (3) found a reduced alkaline phosphatase and commented on its relationship to bone preservation but the isoform targeted was not stated.

Adverse effects of P.mirifica

Numerous different adverse effects were reported (Table 5), the most common events were mastodynia and adverse vaginal problems which are compatible with estrogenic stimulation. Two RCTs having placebo groups (3, 27) totaling 27 participants to compare with *P. mirifica*, and adverse event incidence appeared to be similar in both groups. Chandeying and Sangthawan (25) compared *P. mirifica* with CEE and again adverse event incidences were similar in both groups (not blinded) (Table 5). But they also had 5 hypertriglyceridemic (>400mg/dl) patients treated with *P. mirifica* compared to 1 in the CEE group. The higher baseline TGs for the *P. mirifica* group (121 v 106 mg/dl for CEE), and higher TGs than baselines in other studies suggests these participants were already hypertriglyceridemic at baseline.

Manonai, Chittacharoen (27) measured endometrial thicknesses by ultrasound and performed histology on endometrial biopsies. Thicknesses were unchanged in both PM and placebo groups and biopsies were consistent with atrophic and quiescent endometria. These observations are consistent with the lack of reports of vaginal spotting in these participants.

Of those withdrawn or who withdrew (47 participants), some were reported to be associated with an adverse effect while reasons for withdrawal have insufficient detail. No study reported who decided on participant withdrawal and whether they are blinded. There was no attempt to quantify their severity. Thus, limited data suggests that *P. mirifica* shares the same adverse actions as CEE and similar rates to placebo. But the low number of placebo comparisons, lack of blinding and the wide rate of reporting adverse events, raises concerns. In particular, the propensity for thrombosis was never explored although most studies reported lipid profiles and liver enzymes which remained within normal ranges except the high TGs noted above. However, C-reactive protein was increased 2.7-fold suggesting systemic inflammation (3).

Discussion

This is the first systematic review investigating clinical effect of *P. mirifica* on menopause using a wide range of accepted international bibliographic databases to identify relevant studies and included those in Thailand which provides globally raw *P. mirifica*. Our study adhered to standard methodology of systematic review according to Cochrane and PRISMA checklist (20, 21).

Actions of PM: *P. mirifica* appeared to be estrogenically active. Manonai, Chittacharoen (27) applied objective metrics to detect reduced vaginal pH, vaginal wall thickening and ameliorated the genital tract dysfunction that develop during menopause while bone specific alkaline phosphatase reduced, also accordant with estrogenic actions. Okamura, Sawada (3) using just 12/7 test/control participants also showed robust multiple objective actions comparing baseline and washout while the placebo group remained unaffected. These two studies highlight the importance of including objective, physiological end-points.

Four trials used self-assessed psychological and vasomotor actions as end-points but crucially, none included placebo groups. Placebo-induced amelioration of menopausal symptoms is well-known (30) during (31) and even persisting after placebo treatment ceases (32).

In 24 hormone replacement trials, hot flush intensity in placebo groups were lowered by 58 (CI: 45-68)%, compared to ~75% in test groups (30). This corresponds to our rough calculation of ~78% fall in hot flush severity with *P. mirifica*. A review of 43 phytoestrogen trials found 1-59% reduction of hot flush frequency in placebo groups but no clear additional actions in test groups (2). However, placebo mechanisms can account for 100% of the improved cognitive and psychiatric function (31, 33). A pooled MGCS is biased away from vasomotor symptoms (2/20 domains) towards weaker but entirely psychiatric placebo effects (7 domains). Assessment hot flush *intensity* induces recall error whereas maintaining a diary of hot flush *events* improves objectivity (34), reduces the placebo component (30) and has been validated (35). None of the *P. mirifica* studies used diaries nor hot flush event counting. The magnitude of placebo effects depends on expectation, the nature of the intervention, the context, and interactions with health-care workers especially around the time the questionnaire is completed (36). No *P. mirifica* study provided this information. Thus, the inability to distinguish placebo (and nocebo) effects from genuine physiological actions precludes affirmative conclusions about *P. mirifica* efficacy.

Risk of bias was high for most studies also making these studies inherently unreliable. Furthermore, reporting was lax and the extent of data analysis limited. This has been common in trials on herbal medicines, but transformed by CONSORT now required by many journals.

Molecular targets: Preclinical *in vitro* studies have begun to define the molecular actions of some *P. mirifica* components: principally isoflavones, miroestrols, coumestans at genomic estrogen receptor-alpha(37). However, additional estrogen receptors have emerged including G-protein-coupled estrogen receptor (GPER) to which isoflavones and coumestans bind (38). GPER loss of function accompanies postmenopausal vascular dysfunction (39). ERs may also participate in non-genomic membrane cell signaling, while ER splice variants could play a role in postmenopausal pathology (40, 41). Tissue specific variations in intracellular signaling pathways creates further complexity. Because of the pleotropic and complex estrogenic actions, haphazard preclinical studies to resolve the molecular actions complex mixtures such as *P. mirifica* are unlikely to predict clinical efficacy or safety.

Active ingredients and standardization: The differing *P. mirifica* preparations, inability to chemically define and validate key *P. mirifica* estrogenic components and definition of participants may also affect outcomes. The location (14, 42) (37), plant age (2-4 years), rainfall and growing conditions, processing and storage conditions must be rigorously monitored. The key *P. mirifica* estrogenic components are still unclear. The isoflavones, genestein, daidzein, and their glucosides are commonly used for standardization. However, isoflavone trials have little effect on menopausal symptoms (2), estrogenicity of *P. mirifica* does not correlate with isoflavone content (43), and the isoflavone content in 50 mg of *P. mirifica* is similar in 100 mg of soya, an amount commonly found in human diets (44). The daidzein-C-glycoside, puerarin is often included in standardisations, but its low bioavailability (4nM in plasma after ingesting 10 mg, (45) and only one study showing activity (anti-estrogenic) (46) at <1nMolar likely to exist after 100 mg *P. mirifica*

(containing ~20ug puerarin), makes it probably therapeutically irrelevant. Thus, none of the measured isoflavones are likely to contribute to the *P. mirifica* efficacy and therefore irrelevant for standardisation. Miroestrol (and congeners) are highly potent and thought to confer the greatest ER agonist activity but only one discovered study measured miroestrol (20ug/g) (3) translating to an oral dose of 2ug in the 100mg/day *P. mirifica* powder. Nevertheless, some *P. mirifica* samples yielded 40 ug/100mg dry weight of total miroesterols from powdered *P. mirifica* (16, 47, 48). This compares to typically 600 ug of oral CEE for humans. These small amounts of miroesterols questions the notion of their being the sole active ingredients but instead probably act summatively or synergistically. Coumestrol is also a potent ER ligand (3, 49). Thus, standardization should encompass a key set of estrogens including miroestrols and coumestans in chromatographic 'finger-printing' as required by CONSORT [herbal medicines, (22)] for future clinical trials. The lack of any background pharmacokinetic and metabolic conversion data for *P. mirifica* also complicates analysis of its action. Thus, *P. mirifica* estrogenicity is increased ~9-fold by predigestion with a 'S9' liver extract (37). Most phytoestrogens are rendered more bioavailable and estrogenic by the gut microbiota, commonly by deglycosylation and functional group manipulation (50). How constituents of *P. mirifica* are metabolised is unknown, but vitro studies using native constituents are unlikely to help establishing clinical efficacy or safety. Since the microbiota is highly variable between persons, affects are likely to vary accordingly and reporting mean efficacy is misleading. Alternatively, standardization might rely on bioassay. The simplest and most relevant could be rat vaginal cytology *in vivo* (51). Nevertheless, all the above emphasizes the need for users to titrate doses using a convenient and objective metric such as vaginal pH for optimizing effectiveness.

Safety: While no serious side-effects of oral *P. mirifica* (20-800 mg/day) were noted, reporting was patchy, and each adverse effect was unquantified, apparently without follow-up. Mastodynia and vaginal spotting appeared common but whether they were baseline or placebo effects is unclear. There were no thrombotic measures and no long-term safety assessment of neoplasms. In common with other herbal menopause treatments, *P. mirifica* cannot be said to be safe nor can its efficacy said to outweigh its risks.

The way forward: Many lessons can be learned from the discovered trials for designing future studies. Their underlying theme should be holistic and reflect the traditional use of *P. mirifica* as an anti-aging or 'rejuvenating' agent as recognised by the ancients. As such, this will have the most enduring impact on clinical outcomes and quality-of-life for post-menopausal women and single it out from other treatments. *Firstly*, a retrospective cohort study is needed by recruiting *P. mirifica* users and matched non-users to assess their menopausal and *P. mirifica* dosing history, and current vascular reactivity, metabolic indicators, bone density, vaginal pH, skin elasticity, and cognitive function. These recruits should be prospectively followed up for several years. Besides demonstrating any prolonged benefit, it could demonstrate whether estrogenic treatment initiated before the 'critical period' (ref) during which the irreversible pathologies are set in train. Numerous trials using classical hormone treatments have not resolved this question. *Secondly*, a

short-term intervention study is needed. Its key elements should be: (i) reliable and relevant 'finger printing' and include the relevant parameter in Table 3; (ii) inclusion criteria must account for hormonal status; (iii) dose-efficacy needs establishing using a wide dose range, eg, 0, 10, 50, and 200 mg/day. (iv) a convenient and objective end-point is needed which future users can use for self dose-titration. Self-measured vaginal pH may fulfil this need. (v) the minimal design must have full blinding, a placebo arm, and randomized allocation. (vi) a broad range of end-points is needed accordant with its traditional use of rejuvenation; (vii) accordingly, data needs sub-analyses which needs sufficient power to allow this; (viii) sufficient raw data must be lodged in the public domain to permit alternative analyses (sometimes re-analysis have changed conclusion, most notably in HRT trials). *Thirdly*, long-term outcomes and safety need evaluating. The cohort study should be prospectively followed up over several years. Any predilection of *P. mirifica* to promote gynecological neoplasms and thrombosis remains unanswered. While measuring clotting times and clotting factors, detecting a change in cancer incidences need large and prolonged prospective studies. Nevertheless, robust pharmacovigilance systems for herbal medicines could make major contributions to improving safety of these agents but governments have largely failed to implement these.

Conclusion

Our study suggests that *P. mirifica* could be more potent and efficacious than most other phyto-estrogens for alleviating menopausal symptoms using objective measures. However, no definitive conclusions were drawn from the four trials that tested *P. mirifica* on vasomotor and psychiatric symptoms using the highly subjective MGCS because placebo groups were absent. However, *P. mirifica* may have the potential of maintaining good general health of older women as suggested by its traditional use. Accordingly, there is need for a clear and systematic research strategy that includes (i) cohort (retro- and prospective) studies to define traditional use; (ii) robust independent clinical trials on well characterized patient cohorts, using rigorously applied design criteria and multiple end-points commensurate with multiplicity of pathologies faced by women continuing past the menopause, and (iii) prospective studies and pharmacovigilance.