ALOE VERA AND HEALTH OUTCOMES: AN UMBRELLA REVIEW OF SYSTEMATIC REVIEWS AND META-ANALYSES

Miss Saranrat Sadoyu

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Pharmacy in Clinical Pharmacy
Department of Pharmacy Practice
FACULTY OF PHARMACEUTICAL SCIENCES
Chulalongkorn University
Academic Year 2019

ว่านหางจระเข้และผลสัมฤทธิ์ทางสุขภาพ: การทบทวนแบบครอบคลุมของการทบทวนวรรณกรรมอย่างเป็นระบบและการวิเคราะห์อภิมาน

น.ส.สรัลรัตน์ สโดอยู่

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาเภสัชศาสตรมหาบัณฑิต
สาขาวิชาเภสัชกรรมคลินิก ภาควิชาเภสัชกรรมปฏิบัติ
คณะเภสัชศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2562

Thesis Title	*ALOE VERA* AND HEALTH OUTCOMES: AN UMBRELLA REVIEW OF SYSTEMATIC REVIEWS AND META-ANALYSES
By	Miss Saranrat Sadoyu
Field of Study	Clinical Pharmacy
Thesis Advisor	Assistant Professor Thitima Wattanavijitkul, Ph.D.
Thesis Co Advisor	Professor Nathorn Chaiyakunapruk, Ph.D.

Accepted by the FACULTY OF PHARMACEUTICAL SCIENCES, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science in Pharmacy

..... Dean of the FACULTY OF PHARMACEUTICAL SCIENCES
(Assistant Professor RUNGPETCH SAKULBUMRUNGSIL, Ph.D.)

THESIS COMMITTEE

..... Chairman
(Associate Professor WANCHAI TREYAPRASERT, Ph.D.)

..... Thesis Advisor
(Assistant Professor Thitima Wattanavijitkul, Ph.D.)

..... Thesis Co-Advisor
(Professor Nathorn Chaiyakunapruk, Ph.D.)

..... Examiner
(Assistant Professor NUTTHADA AREEPIUM, Ph.D.)

..... External Examiner
(Associate Professor Surasak Saokaew, Ph.D.)

สรัลรัตน์ สโดอยู่ : ว่านหางจระเข้และผลสัมฤทธิ์ทางสุขภาพ: การทบทวนแบบครอบคลุมของการทบทวนวรรณกรรมอย่างเป็นระบบและการวิเคราะห์อภิมาน. (*ALOE VERA* AND HEALTH OUTCOMES: AN UMBRELLA REVIEW OF SYSTEMATIC REVIEWS AND META-ANALYSES) อ.ที่ปรึกษาหลัก : ผศ. ภญ. ดร.ธิติมา วัฒนวิจิตรกุล, อ.ที่ปรึกษาร่วม : ศ. ภก. ดร.ณธร ชัยญาคุณาพฤกษ์

การทบทวนแบบครอบคลุมนี้มีวัตถุประสงค์เพื่อรวบรวมและประเมินผลสัมฤทธิ์ทางสุขภาพของการใช้ว่านหางจระเข้ วิธีการศึกษา คัดเลือกงานวิจัยที่ศึกษาผลของการใช้ว่านหางจระเข้ต่อผลสัมฤทธิ์ทางสุขภาพในรูปแบบการทบทวนอย่างเป็นระบบและวิเคราะห์อภิมานของการทดลองทางคลินิก สืบค้นจากฐานข้อมูล PubMed, Scopus, EMBASE, Cochrane database of systematic reviews, CINAHL, และ AMED จนถึงเดือน ตุลาคม พ.ศ. 2562 โดยผู้วิจัยสองคนคัดเลือก สกัดข้อมูลและประเมินคุณภาพของงานวิจัยอย่างเป็นอิสระต่อกัน จากนั้นจัดระดับความน่าเชื่อถือของหลักฐานโดยแบ่งออกเป็นความน่าเชื่อถือระดับสูงมาก, ระดับสูง, ระดับแนะนำ, ระดับต่ำ และไม่มีนัยสำคัญ ผลการศึกษาพบว่ามีงานวิจัยผ่านเกณฑ์การคัดเลือกจำนวน 10 การศึกษา รายงานผลของการใช้ว่านหางจระเข้ในผลสัมฤทธิ์ทางสุขภาพ 71 ข้อ เมื่อทดสอบด้วย Random effects model พบว่า 47 ข้อ (ร้อยละ 67) มีนัยสำคัญทางสถิติ ($p \leq 0.05$) แต่มีเพียง 3 ข้อที่มีระดับความน่าเชื่อถือสูง ได้แก่ ประโยชน์ของว่านหางจระเข้ในการป้องกันการเกิดหลอดเลือดดำอักเสบที่เกิดจากการหยดยาทางหลอดเลือดดำในความรุนแรงระดับสอง (RR: 0.18, 95% CI: 0.10-0.32) และการป้องกันหลอดเลือดดำอักเสบจากยาเคมีบำบัดทั้งการเกิดหลอดเลือดดำอักเสบในความรุนแรงระดับสองและความรุนแรงรวมทุกระดับ (OR: 0.10, 95% CI: 0.07-0.14 และ OR 0.13, 95% CI: 0.08-0.20 ตามลำดับ) โดยสรุปงานวิจัยนี้ยืนยันว่าว่านหางจระเข้สามารถใช้เพื่อป้องกันภาวะหลอดเลือดดำอักเสบจากการให้ยาและยาเคมีบำบัดทางหลอดเลือดดำโดยเฉพาะในระดับที่มีความรุนแรงมาก ส่วนข้อบ่งใช้อื่นพบว่าว่านหางจระเข้มีประสิทธิภาพเช่นกัน แต่การศึกษาส่วนใหญ่มีข้อจำกัด ได้แก่ จำนวนผู้เข้าร่วมการศึกษาน้อยและมีคุณภาพของระเบียบวิธีวิจัยต่ำ

สาขาวิชา	เภสัชกรรมคลินิก	ลายมือชื่อนิสิต .....
ปีการศึกษา	2562	ลายมือชื่อ อ.ที่ปรึกษาหลัก .....
		ลายมือชื่อ อ.ที่ปรึกษาร่วม .....

6176113633 : MAJOR CLINICAL PHARMACY
KEYWORD: aloe vera, health outcome, umbrella review, credibility assessment

Saranrat Sadoyu : *ALOE VERA* AND HEALTH OUTCOMES: AN UMBRELLA REVIEW OF SYSTEMATIC REVIEWS AND META-ANALYSES. Advisor: Asst. Prof. Thitima Wattanavijitkul, Ph.D. Co-advisor: Prof. Nathorn Chaiyakunapruk, Ph.D.

This umbrella review aims to summarize and assess the effects of *Aloe vera* on health outcomes. Methods: Only systematic reviews and meta-analyses of clinical trials that investigated the effects of *Aloe vera* on health outcomes were eligible. PubMed, Scopus, Embase, Cochrane database of systematic reviews, CINAHL, and AMED were searched from inception to October 2019. Two independent reviewers extracted data, assessed the methodological quality, and rated the credibility of evidence according to established criteria into convincing, highly suggestive, suggestive, weak, and not significant. Results: Ten articles reporting 71 unique outcomes of *Aloe vera* were included. Of these, 47 (67%) were nominally statistically significant based on the random-effects model ($p \leq 0.05$). Only 3 outcomes were supported by highly suggestive evidence including the benefits of *Aloe vera* in the prevention of second-degree infusion phlebitis (RR: 0.18, 95% CI: 0.10-0.32) and chemotherapy-induced phlebitis based on the second-degree of severity and overall incidence (OR: 0.10, 95% CI: 0.07-0.14, and OR: 0.13, 95% CI: 0.08-0.20, respectively). Conclusions: The current evidence suggests the benefits of *Aloe vera* in the prevention of phlebitis induced by chemotherapy and intravenous infusion, particularly in the severe stage. *Aloe vera* also showed favorable effects in other indications, but the majority of the evidence had limitations including small sample size and poor methodological quality.

Field of Study:	Clinical Pharmacy	Student's Signature .....
Academic Year:	2019	Advisor's Signature .....
		Co-advisor's Signature .....

ACKNOWLEDGEMENTS

My fully thanks and appreciation go to Asst. Prof. Thitima Wattanavijitkul and Prof. Nathorn Chaiyakunapruk, my advisor and co-advisor, for their help, understanding, wisdom, patience, and encouragement and for pushing me farther than I thought I could go.

I would like to express my sincere gratitude to Miss Chidchanok Rungruang. This thesis can be completed because of her kind support and help as a second reviewer. I would also like to thank Asst.Prof. Peerawat Jinatongthai for his help on databases searching and search strategy suggestions.

In addition, I would like to thank Asst. Prof. Ratree Sawangjit and Prof. Ammarin Thakkinstian fort their advice and valuable information. My thanks and appreciations also go to Mr. Sajesh K Veettil and Mr. Krittin Bunditanukul, for sharing their knowledge and technical know-how in statistical analysis.

I would also like to thank my beloved friends Natnicha Chirapreeya, Arissara Karnjananggulpan, Nalinee Kruetiwa, and Mathipot Srisarakam who always be there and really helped me cope with the problems that I faced. Thanks to Chanatinat Anamnart for translating from Chinese to the English language. Special mention and extremely thankful goes to Jam whose materials, advice, and encouragement have been priceless. Also, my classmates, Jip, Beil, Ple, Imm, Little Imm, Wan, and Bew, thank you for being around and study together.

Finally, I would like to acknowledge the support and love of my family – my parents and my sisters. This thesis would not have been possible without them.

Saranrat Sadoyu

TABLE OF CONTENTS

Page

List of Tables

List of Figures

CHAPTER I

Introduction

1. Background and rationale

The use of complementary and alternative medicine for health maintenance and disease management has been growing rapidly worldwide. The national survey of the general population in the United States and European countries reported that 40% (1) and 25.9% used complementary and alternative medicine as alternative therapy in the past 12 months (2). In Thailand, one previous study found a high prevalence of herbal medicine use among 35.9% of chronic disease patients (3).

Aloe vera (synonym: *Aloe barbadensis* Miller), one of more than 400 species belonging to the Liliaceae family, has been used in traditional medicine for centuries. It is a succulent cactus-like plant originated in Africa and commonly grown in hot and dry climates. Numerous medicinal products are made from *Aloe vera* leaf are widely available, contained active constituents such as vitamins, enzymes, minerals, sugars, and amino acids which are associated with various pharmacological activities including fungicidal, antiviral, antibacterial, anti-inflammatory, laxative, immunomodulating, and anticancer effects (4, 5). In the 2012, *Aloe vera* was ranked the 20th among best-selling dietary supplements in the United states (6). In Europe and United states, oral administrations of *Aloe vera* gel extract and latex are indicated for constipation and food supplements (7, 8). Additionally, WHO published a monograph of *Aloe vera* gel summarizing the topical use of *Aloe vera* for the external treatment of minor wounds and inflammatory skin disorders (9). However, its efficacy is mostly based on historical use or scientific theory rather than clinical evidence (6, 9).

The benefits of *Aloe vera* have been examined widely through both animal studies and clinical trials, resulting in the growing number of systematic reviews and

meta-analyses with attempt to summarize and clarify these findings. Previously published meta-analyses demonstrated that *Aloe vera* was effective in several indications; for example, reduction of time to complete wound healing (10), improving symptoms of irritate bowel syndrome (11). Interestingly, recent meta-analyses reported potential benefits of *Aloe vera* in blood glucose reduction (12, 13), prevention and treatment of phlebitis (14). Unfortunately, although numerous meta-analyses have been published, each of them often focused only on one particular indication, which made it difficult for clinicians or healthcare providers to synthesize the overall benefits of *Aloe vera*.

To date, there has been no effort to integrate previously published systematic reviews and meta-analyses across many indications of *Aloe vera* and determine quality of the evidence which can be done by performing an umbrella review; a new approach currently recognized as one of the highest levels of evidence synthesis methods (15-17). Umbrella review consists in the meta-analyses following a uniform approach for all factors to allow their comparison, systematic review of previous systematic review, and also assess whether there is evidence for biases. Umbrella review has been increasingly conducted and published in this decade to address an extensive and high-quality evidence base around a topic (18-25). This method will help summarize the evidence from multiple systematic reviews and meta-analyses, assessing the strength of evidence and extent of possible biases to ascertain the confidence in its benefits (15, 16).

Therefore, this study aimed to conduct the umbrella review to systematically summarize and assess the effects of *Aloe vera* use on various types of health outcomes for determining its benefits in each indication from previously published systematic reviews and meta-analyses to provide the wide picture for real-world clinical practice.

2. Research question

What are the effects of *Aloe vera* on health outcomes, and how credible is the evidence?

3. Objective of the Study

To summarize existing systematic reviews and meta-analyses that assessed the effects of *Aloe vera* on health outcomes.

4. Scope of the Study

Systematic reviews and meta-analyses of clinical trials comparing *Aloe vera* with any comparators and reported health outcomes.

5. Operational definitions

Aloe vera is a plant belonging to the Liliaceae family, of which also known as *Aloe barbadensis* Mill. *Aloe vera* used in this umbrella review refers to either orally or topically used of the crude *Aloe vera* components from any part of this plant (for example, leaf, fresh juice, fresh gel, and fresh stem) and the *Aloe vera*-derived products such as *Aloe vera*-extracted powder, *Aloe vera* gel, *Aloe vera* cream, *Aloe vera* ointment, *Aloe vera* mucilage, and *Aloe vera* mouthwash.

Health outcomes are defined as any changes in health that result from treatments or interventions including;

- Clinical outcomes: efficacy on indication of *Aloe vera* (e.g. pain reduction in irritate bowel syndrome, reduction in fasting plasma glucose in diabetes, time to complete wound healing)
- Patient reported outcomes (e.g. quality of life, satisfaction)

Systematic review is a type of study that conducted in an attempt to compile all existing evidence that fits pre-specified eligibility criteria to answer a specific research question. The key characteristics are as follows:

- a clearly stated set of objectives with an explicit, reproducible methodology;
- a systematic search that attempts to identify all studies that would meet the eligibility criteria;

Meta-analysis is defined as a study that use statistical techniques to integrate and summarize results of included studies.

Duplicate article is defined as an article that published the similar material more than once, by the similar author or publisher.

6. Expected benefits of the study

The information resulted from this study can influence clinicians and healthcare providers's concept of the quantity and quality of existing evidence of *Aloe vera* effectiveness in various health outcomes.

CHAPTER II

LITERATURE REVIEW

1. *Aloe vera*

Aloe vera (synonym: *Aloe barbadensis* Miller), one of more than 400 species of Aloe, belongs to the Liliaceae (Aloeaceae) family (4, 5). *Aloe vera* is the most popular and widely used species (26) among the genus Aloe such as *Aloe vera, Aloe barbadensis, Aloe ferox, Aloe chinensis* (27). It is also known as aloe, burn plant, lily of the desert, and elephant's gall (27).

1.1 Botanical data

Aloe vera is a stem less or very short-stemmed cactus-like plant originated in Africa and commonly grown in hot and dry climates. It probably native to north Africa and has been subsequently introduced in the Mediterranean region, and most of the tropical areas worldwide such as Asia, the Bahamas, Central America, Mexico, and the southern region of the United States (28).

Botanical name: *Aloe vera* is the synonym of the species name '*Aloe barbadensis* Mill' regarding most books and formularies.

Family: Liliaceae (Aloeaceae)

Genus: Aloe

1.2 Phytochemical Composition

Aloe vera plant contains the high-water content in the main feature, ranging from 99–99.5% while the 0.5–1.0% remaining is the solid material which contains plenty of potentially active compounds, including vitamins, minerals, enzymes, simple and complex polysaccharides, phenolic compounds, and organic acids (26). The leaves of *Aloe vera* are green and fleshy, has been found in the height range from a few centimeters to over 2 meters. *Aloe vera* leaf can be divided into three layers including;

- The outer layer, also known as the outer green epidermis, is a thick cuticle called rind and weighs around 20-30% of the whole plant leaf. This layer has been reported the containing of anthraquinones and glycosides (9, 29).
- The middle is a thin mucilaginous layer occurring below and next to the outer thick green rind. The appearance of this layer is a yellow latex or exudate (commonly called as aloe sap). This latex mainly contains phenolic compounds such as anthraquinones, coumarins, anthrones, chromones, pyrones, and flavonoids (30).
- The inner leaf pulp called Aloe gel, a synonym to the inner leaf, inner leaf fillet, or fillet, lies in the center of the leaf. The Aloe gel predominantly contains water more than 90%, while the remaining contains more than 75 known substances including polysaccharides (i.e. pectins), hemicelluloses, glucomannan, acemannan, and mannose derivatives), vitamins (i.e. A, B, C, and E), calcium, lipids, sterols, amino acids, and enzymes (9).

The physical and chemical constituents in the *Aloe vera*-derived products can be differed depending on several factors. For example, the climate conditions while harvesting and storage, the source (e.g. part of the plant used), seasonal and grower influences, and processing techniques (31).

1.3 Pharmacological effects

1.3.1. Wound healing: *Aloe vera* has long been used for wounds treatment more than hundred years. Mannose 6-phosphate, a primarily polysaccharide found in *Aloe vera* gel, has been considered as the active ingredient of wound healing by directly stimulating the activity of macrophages, increasing amounts of collagen,

proteoglycan synthesis, and accelerates the healing process (26, 31). Acemannan, another polysaccharide that commonly found in *Aloe vera*, has been reported the effects in the increasing of periodontal ligament cell proliferation, collagen and alkaline phosphatase activity, and enhancing the growth factor upregulation in primary human periodontal ligament cells (32). Glycoproteins and saponin were also reported to have wound-healing activity (33, 34). The wound infection has been reported to be prevented by the antibacterial properties of anthraquinones (31).

Previous studies in animal models suggested the wound healing effect of *Aloe vera* (35, 36). In clinical trials, *Aloe vera* cream demonstrated a faster rate of healing and epithelialization than silver sulfadiazine cream in the site treated for second-degree burn wounds (37). Application of *Aloe vera* gel has also been reported to demonstrate shorter average healing times than the petroleum jelly gauze (38) and silver sulfadiazine cream (39), and accelerate the healing of split-thickness skin graft donor sites (40). Nevertheless, while most clinical trials reported the potential benefits of *Aloe vera*, some suggested that *Aloe vera* might slow the wound healing rate or no statistical difference when compared with placebo (41, 42).

1.3.2. Antidiabetic and anti-hyperlipidemic Activity: Polysaccharide possesses a major role in hypoglycemic activity by increasing insulin level (28). Phytosterols derived from aloe gel, such as ophenol (Lo) and cycloartanol (Cy), altered the expressions of genes resulted in glucose level reduction and ameliorated obesity-associated metabolic disorders in Zucker diabetic fatty rats (43). Aloe-emodin-8-O-glycoside has been reported to increase glucose uptake and transformed it into glycogen. (44). Besides, some investigators proposed that *Aloe vera* improved insulin sensitivity and reduced body fat through adenosine monophosphate-activated muscle protein kinase activation (45).

Several studies in rodents suggested the anti-hyperglycemic and anti-hyperlipidemic effect of *Aloe vera* (45-47). Furthermore, previously published clinical trials reported the efficacy of *Aloe vera* among type 2 diabetic patients. Given its juice, one tablespoonful twice daily for at least 2 weeks, the effects in the reduction of fasting blood glucose were found, but cholesterol levels were not affected (48, 49), Other clinical trials found that 300 mg capsule derived from *Aloe vera* gel given twice daily for 2 months not only lowered the fasting blood glucose (FBG) and glycosylated hemoglobin A1c (HbA1c) significantly but also lowered total cholesterol and low-density lipoprotein cholesterol (LDL) levels (52). Choudhary et al (53) reported that taking 100 mg and 200 mg of Aloe vera gel powder for 3 months in diabetic patients, resulting in the reduction of postprandial glucose, FBS, total cholesterol, very-low-density lipoprotein cholesterol (VLDL) and LDL-level, and increasing high-density lipoprotein cholesterol (HDL) level. In those with prediabetes or early diabetes, *Aloe vera* capsule consumption also lowered FBG, HbA1C, insulin level, body weight, and fat mass (34, 50).

1.3.3. Laxative effect: Anthraquinones, presented in the latex of *Aloe vera*, is a potent stimulant laxative. It help stimulate mucous secretion and increases intestinal water content and peristalsis (51). *Aloe vera* laxative preparations have been approved for use in constipation treatment (7). However, the U.S. Food and Drug Administration has no longer approved *Aloe vera* latex as an over-the-counter drug for constipation treatment to date because of a lack of sufficient data to prove its safety (52).

1.3.4. Effect on gastric acid secretion and ulcers: The lectins in *Aloe vera* inhibit acid producing cells, resulting in the inhibition of gastric acid output (27, 53). Polysaccharides and anthraquinones in *Aloe vera* also help reduce peptic ulcers by controlling gastric secretion. *Aloe vera* gel has been reported its antibacterial

properties in both susceptible and resistant *H. pylori* strains. Additionally, antiulcer effect of Aloe vera has been reported. *Aloe vera* decreased mean ulcer index in rats more than the omeprazole-treated group (10 ± 1.96 and 20 ± 1.79 respectively, $p < 0.001$) (54). Another study in humans reported that the combination of *Aloe vera* and sucralfate can reduce gastric inflammation and ulcer sizes, increase epithelial cell proliferation, and lengthene gastric glands (55).

1.3.5. Anti-inflammatory action: Efficacy of *Aloe vera* has been reported in the cyclooxygenase pathway, thromboxane, and prostaglandin inhibition, resulting in pain reduction and acceleration of the healing process (56). The previous study in rats has also reported the effect of *Aloe vera* gel in acute-inflammatory reduction through its plant sterol called Lupeol (57).

1.3.6. Antimicrobial Activity: *Aloe* vera showed efficacy in inhibiting *Staphylococcus aureus* growth when high concentrations were used, while moderate concentrations showed efficacy in inhibiting *Escherichia coli, Salmonella typhi,* and *Pseudomonas aeruginosa* (58, 59). Besides, *Aloe vera* gel showed a bactericidal effect and prevent *Pseudomonas aeruginosa* from adhering to human lung epithelial cells (60).

Moreover, several antiseptic agents including salicylic acid, cinnamonic acid, urea nitrogen, lupeol, phenols, and sulfur found in *Aloe vera* have been reported an inhibitory effect on bacteria, viruses, and fungi (4). The lectin-containing fraction of *Aloe vera* gel demonstrates the effect in inhibition of Cytomegalovirus growth, by interfering the protein synthesis (61). Additionally, the anthraquinone derivatives (e.g. Aloe-emodin and chrysophanol) exhibited antiviral activity on enveloped viruses such as influenza, varicella zoster, and herpes simplex viruses through the reduction of virus-induced cytopathic effect and the inhibition of its replication (62, 63).

1.3.7. Dentistry: The role of *Aloe vera* in dentistry is applied because of its antibacterial and wound-healing effects. *Aloe vera* has been reported potential benefits in gum disease treatment (e.g. gingivitis and periodontitis) (64). Previous studies reported the potential effect of *Aloe vera* in quality of life and symptom improvement, and pain reduction in various types of dentistry problems including oral lichen planus (65), aphthous ulcer (66), and oral submucous fibrosis (67).

1.3.8. Anticancer activity: The polysaccharide fraction of *Aloe vera* gel help prevent the formation of benzo[α]pyrene- DNA adducts and stimulate the immune response which resulting in its chemo-preventative effects (31, 53). Aloe-emodin, an anthraquinone derived from *Aloe vera*, has been reported anti-neoplastic effects by inhibited the malignant cells' growth (68).

1.3.9. Skin protection and hydration: *Aloe vera* help increase hydration and the penetrating ability of the skin resulting in the reduction of flaky scalp and skin (69). Besides, G1C2F1, a small-molecular-weight immunomodulator in *Aloe vera* gel, helps prevent ultraviolet B (UVB)-induced damage in the epidermal Langerhans cells (LC) by preventing the UVB-induced immune suppression (70). *Aloe vera* has also been reported potential effects in improving impaired skin integrity, decrease fine wrinkle, and erythema, resulting in the improvement of irritant contact dermatitis and dry skin (68).

1.4 Therapeutic use

Aloe vera has long been used worldwide as a traditional medicine such as in Greece, Egypt, India, China, and Japan for more than 2000 years (26). It is used in various ways, as over-the-counter (OTC) medications, as self-care or home remedies. However, its efficacy is mostly based on historical use or scientific theory rather than clinical evidence. Generally, *Aloe vera* gel has been used widely to soothe wounds, burns, skin irritations, and inflammatory skin disorders by an external application. The

latex of *Aloe vera* has been used due to its cathartic effects. The treatment of acne, hemorrhoids, psoriasis, anemia, glaucoma, tuberculosis, seborrheic dermatitis, and fungal infections by *Aloe vera* gel has also been described in folk medicine (9, 29).

1.5 Dosage and administration

Constipation: Dried latex 40-110 mg/day or 100 mg as a single dose in the evening, corresponding to 10-30 mg of hydroxyanthraquinones (9) for adults and children aged more than 10 years. The maximum dose of hydroxyanthracene glycosides in *Aloe vera*-derived preparation is 30 mg/day (7).

Diabetes (type 2): 5–15 mL of *Aloe vera* juice 2 times/day (51).

Wound healing: No consensus exists regarding the dosage and administration of topical *Aloe vera* for wound healing. The results from previous literatures indicated the potentially benefit of 0.5% *Aloe vera* (gel or cream) applies on affected area twice daily (71).

Genital herpes and psoriasis vulgaris: Hydrophilic cream of 0.5% (by weight) of a 50% ethanol extract of *Aloe vera*, apply on affected area 3 times/day for 5 consecutive days per week has been used for up to 2 weeks in treatment of genital herpes and up to 4 weeks in psoriasis vulgaris (51).

1.6 Contraindications

- Allergic to plants in the Liliaceae family (i.e. onions and garlic).
- Intestinal obstructions and stenosis, atony, appendicitis, abdominal pain of unknown origin (9).
- Oral administration of *Aloe vera* is not recommended in children (age < 10 years), pregnancy, and breastfeeding mothers (9, 29).

1.7 Warning and precautions

- Avoid the long-term use of *Aloe vera* as a stimulant laxatives due to its potential to cause intestinal dysfunction and laxative dependence (9).

- Avoid using in patients with kidney disorders due to its potential to cause electrolyte imbalance. Hypokalemia caused by *Aloe vera* has also been reported (29).

1.8 Drug interactions

- Cytochrome P450 substrates: *In vitro*, *Aloe vera* juice was found to inhibit CYP3A4 and CYP2D6 (72).
- Sevoflurane: Concomitant use of *Aloe vera* and sevoflurane may cause excessive perioperative bleeding based on additive antiplatelet effects of these 2 agents (73).

1.9 Stability

The amounts of active ingredients can vary among *Aloe vera* preparations, depending on harvesting and storage conditions, part of plants used, the time of used after harvesting, and extraction methods (31, 53). The clear mucilaginous gel of *Aloe vera* was found stable when pasteurized at 75–80°C for less than 3 minutes. However, longer times and higher temperatures may alter the chemical composition of the gel (9, 74).

Chemically preserved fresh *Aloe vera* gel incubated at 40 °C or stored at room temperature for 48 hours showed the rheological properties degradation and a reduction of polysaccharides content and composition (75). The rapid deterioration of Aloin, which contained in whole leaf extract of *Aloe vera*, was also detected during storage, especially at higher temperatures (76).

1.10 Adverse reaction

Dermatologic: Minor adverse effects were reported in clinical trial using the topical preparation of *Aloe vera.* For example, itching and slightly tingling sensation

(29). However, papular dermatitis and generalized eczematous were reported by male patient after using the oral and topical *Aloe vera* gel (77).

Gastrointestinal: Prolonged ingestion of *Aloe vera* was associated with abdominal pain, vomiting, diarrhea, pseudomelanosis coli, cathartic colon, and increase risk of colorectal cancer (29, 51). Hepatotoxicity was also found as case reports (78, 79).

Carcinogenic: *Aloe vera* whole leaf extract showed evidence of carcinogenic activity in rats, classified by the International Agency for Research on Cancer as a possible human carcinogen (Group 2B), along with other natural products such as *Ginkgo biloba* extract and kava extract (29).

Renal: Prolonged ingestion of *Aloe vera* latex may cause potassium depletion and electrolyte imbalance because of its laxative effect (4).

2. Health outcomes

Health was defined by the Constitution of the World Health Organization in 1948 as ‘a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity’ (80). The novelty of widening the concept of health into a multidimensional construct to include psychological and social dimensions of people's lives as well as focusing positively on well-being was significant (81).

Health outcomes are important for assessing effectiveness of medical interventions and determining results in clinical practice because outcome measure is one of the indicators generally used to evaluate the quality of interventions or working processes (82). Although no standard definition of health outcomes used to date, the definition of health outcomes has been mentioned by various literature and organization as follows;

- The International Consortium for Health Outcomes Measurement (ICHOM), a non-profit organization founded in 2012, defines health outcomes simply as 'the results of treatment that patients care about most'. This organization has defined and developed more than 12 standard datasets for outcomes measurement which increasingly used worldwide (83).
- According to Canadian institute for health information, health outcomes are changes in health that result from measures or specific health care investments or interventions. Patient-reported outcomes are essential to understanding whether health care services and procedures make a difference to patients' health status and quality of life (84).
- In Australia, Australian Health Outcomes Collaboration defines a health outcome as 'health outcome is a change in the health of an individual, or a group of people or population, which is wholly or partially attributable to an intervention or series of interventions'. Health outcome measures can include clinical/biomedical indicators, indicators related to survival, performance indicators, standardized clinical assessments, and patient-reported outcome measures (85).

Health outcomes also defined as those events occurring as a result of an intervention. These may be measured clinically (physical examination, laboratory testing, imaging), self-reported, or observed (such as gait or movement fluctuations seen by a healthcare provider or caregiver). Some health outcomes require complex assessments to determine if they are present or absent. There is a range of standardized and validated measures of health status/health-related quality of life

(HRQoL) that along with clinical and performance indicators can be used to assess the outcomes of treatment interventions (86).

In conclusion, health outcomes are the changes in health that result from treatments or interventions. These differences in health status can be measured in various ways such as patient self-report, clinical assessment, and observation by the healthcare provider or caregiver. Measuring health outcomes can help healthcare workers making decisions about providing care for their patients.

3. Systematic review and meta-analysis

A systematic review involving the identification, selection, appraisal, and summary of primary literature addressing a focused clinical question using methods to reduce the likelihood of bias (87, 88). A systematic review is one of a medical subject headings (MeSH) term which defined by the National Library of Medicine as 'A review of primary literature in health and health policy that attempts to identify, appraise, and synthesize all the empirical evidence that meets specified eligibility criteria to answer a given research question. Its conduct uses explicit methods aimed at minimizing bias in order to produce more reliable findings regarding the effects of interventions for prevention, treatment, and rehabilitation that can be used to inform decision making'.

Grant et al (89) described the systematic review as the best-known type of review which seeks to systematically search for, appraise, and synthesis research evidence, often adhering to the guidelines on the conduct of a review provided by the Cochrane Collaboration or the NHS Centre for Reviews and Dissemination. It is transparent in the reporting of its methods to facilitate others to replicate the process.

Systematic review described in the Preferred Reporting Items for Systematic reviews (PRISMA) Statement is an attempt to collate all empirical evidence that fits

pre-specified eligibility criteria to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing reliable findings from which conclusions can be drawn and decisions made (90).

The key characteristics of a systematic review are:

- a clearly stated set of objectives with an explicit, reproducible methodology;
- a systematic search that attempts to identify all studies that would meet the eligibility criteria;
- an assessment of the validity of the findings of the included studies, for example through the assessment of risk of bias;
- systematic presentation, and synthesis, of the characteristics and findings of the included studies.

A meta-analysis is a type of study that uses statistical methods to perform a quantitative synthesis of existing data. Meta-analysis can help increases the precision of effects compared to each trial and provide more generalizability of study findings (17). According to PRISMA statement, meta-analysis is the use of statistical techniques to integrate and summarize the results of included studies. Many systematic reviews contain meta-analyses, but not all. By combining information from all relevant studies, meta-analyses can provide more precise estimates of the effects of health care than those derived from the individual studies included within a review (90).

4. Previous systematic reviews and meta-analyses of *Aloe vera*

4.1 Wound healing

Maenthaisong et al, (10) conducted a systemic review and meta-analysis from 4 clinical trials investigating effects of *Aloe vera* used in burn wound healing. They reported that the mean difference of healing time in the *Aloe vera* group was

significantly shorter than the control group (8.79 days, p = 0.006). However, the differences in products and outcome measures in their study make the authors cannot suggest a specific conclusion regarding the effect of *Aloe vera* on burn wound healing.

According to Dat et al. (71), they reported in their systematic review that only a small number of high-level evidence supported the use of *Aloe vera* topical agents in acute and chronic wound treatment. However, the authors could not perform a meta-analysis in their review because replication of trials with the same wound type and intervention was lacking.

In conclusion, Although *Aloe vera* showed a potential benefit in wound healing, there is currently an absence of high-quality clinical evidence supported its use in acute and chronic wound treatment. The difference in results might cause by the difference in study design, and products variation of *Aloe vera*.

4.2 Antidiabetes and antilipidemia

Zhang et al (91) performed a systematic review and meta-analysis from 5 RCTs included evaluating the effects of *Aloe vera* in prediabetes and early non-treated diabetes mellitus. They reported a significant reduction of FBS, HbA1c, total cholesterol, LDL, TG, and increased HDL level in *Aloe vera* receiving group, but results are inconclusive because of small and poor quality of RCTs included. Among prediabetic and diabetic patients, results from a meta-analysis conducted by Dick et al (92) suggested benefits of oral administration of *Aloe vera* in FBG and HbA1c reduction, similar to Zhang et al (91). In contrast to Suksomboon et al (93), which reported some potential benefits of *Aloe vera* in FBS reduction, but no effect on HbA1c reduction was found. Despite the useful information provided by these three studies, implementation into clinical practice should be done cautiously and additional investigations are needed due to noticeable heterogeneity, limited

number, and poor quality of primary studies included in systematic reviews and meta-analyses.

4.3 Gastrointestinal disease

Aloe vera showed a potential benefit in patients with inflammatory bowel syndrome (IBS) from previously published 1 systematic review and 1 meta-analysis. Langhorst et al (94), performed a systematic review to identify the effectiveness of complementary and alternative medicines in IBS. *Aloe vera* gel, mentioned as one of the useful herbal medicines, showed significant improvements in clinical signs and quality of life in patients with active ulcerative colitis. Hong et al (11) published a meta-analysis of 3 RCTs with a total of 151 patients that indicated the benefits of Aloe vera in the improvement IBS symptoms and response rate. However, no statistically significant differences were found. The authors discussed that might due to the small number of participants in each study, and further, summarizing and grading on this evidence are required.

4.4 Dentistry

Numerous systematic reviews and meta-analyses were performed to determine the effectiveness of *Aloe vera* in dentistry. Both Nair et al (95) and Mangaiyarkarasi et al (96), reported the benefit of *Aloe vera* in aphthous stomatitis, oral submucous fibrosis, oral lichen planus, gingivitis, and radiation-induced oral mucositis in their systematic reviews which included more than 10 RCTs. Al-Maweri et al (97) demonstrated a promising effect of *Aloe vera* in clinical improvement and pain reduction among patients with oral submucous fibrosis by meta-analysis consists of 6 RCTs, however, statistically significant differences in pain reduction between *Aloe vera* and control group was reported only at the end of the first and second month, but not found in the third month. The researchers stated that small sample

sizes and high risk of bias in some of the included studies might affect the credibility of their results.

A meta-analysis evaluated the effectiveness of *Aloe vera* in oral lichen planus conducted by Ali and Wahbi (98) suggested that topical *Aloe vera* showed more benefit than placebo on clinical improvement and pain reduction, but only comparable to triamcinolone acetonide. However, findings from their meta-analysis included data from only 3 RCTs, and high heterogeneity was found (I^2 = 94%. Dhingra et al (99), including 2 RCTs in gingivitis patients in his systematic review, reported that *Aloe vera* dentifrices showed potential efficacy in reducing plaque and gingival inflammation was similar to control group.

Therefore, even though many systematic reviews and meta-analyses have been performed, the summarize of the effectiveness of *Aloe vera* in dentistry and the evaluation of the quality of this evidence is needed.

4.5 Phlebitis

Zheng et al (14) reported the potential benefits of *Aloe vera* external application in preventing or treating infusion phlebitis compared with no intervention or $MgSO_4$. *Aloe vera* was found to prevent infusion phlebitis, however, no clear evidence supported the effect of *Aloe vera* in the treatment of infusion phlebitis when compared with other interventions in this study, such as 75% alcohol and 75% $MgSO_4$. The authors suggested that their findings were limited by risk of selective outcome reporting and the poor methodological quality of the included studies.

In conclusion, findings from previously systematic reviews and meta-analyses suggested potential benefits of *Aloe vera* in various health outcomes. However, inconclusive results were found in some outcomes. Therefore, combining and appraising the numerous pieces of evidence is essential to provide a wider picture of *Aloe vera* on health outcomes.

5. Umbrella review

Systematic review and meta-analysis are important research design aim to synthesize the findings and assess the biases of existing pieces of evidence. However, results from several systematic reviews and meta-analyses might be inconclusive because of several types of biases. Moreover, a single meta-analysis of a treatment comparison for a single outcome might offer a limited view if there are many treatments or many important outcomes to consider (16). In recent years, systematic reviews and meta-analyses have been published in various fields gradually. Bastian et al (100), reported that 11 systematic reviews were being published per day. These too many reviews can be the problems for clinicians and healthcare providers in sorting evidence and synthesis the findings to inform their questions and making decisions in real-world clinical practice.

To summarize the numerous findings from existing systematic reviews and meta-analyses, umbrella review has been introduced. Ioannidis et al (16) defined umbrella reviews as a systematic review that assembles several systematic reviews on the same disease or condition and also combines statistical results from several meta-analyses as appropriate. Biondi-Zoccai (17) stated that umbrella review is a synthesis of systematic reviews, only considers the inclusion of the higher level of evidence including systematic reviews and meta-analysis, which allows the results of relevant assessments in a review question can be compared and contrasted.

According to Joanna Briggs Institute (101), the term umbrella review is a systematic review that draws together evidence from a series of other systematic reviews. Umbrella review help provide an overview of research within a specific area. In addition, Fusar-Poli and Radua (15) described an umbrella review as one of the highest levels of evidence synthesis currently available, as shown in figure 1. This umbrella review approach can be used to summarize previously published

systematic reviews or meta-analyses and also provide the repetition of the meta-analyses to allow the comparison.

In conclusion, umbrella review can be described as a new evidence-based synthesis method that aims to combine and compare findings, determine statistical data, and assess the extent of possible biases and strength of evidence from existing systematic reviews and meta-analyses.

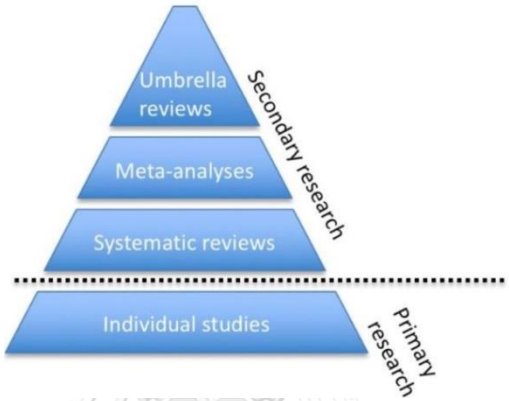

Figure 1 Ranking of evidence synthesis methods

5.1 Previously published umbrella reviews

Umbrella review has been conducted and published increasingly in this past decade. Furthermore, the growing number of protocols of umbrella reviews have been recently published. Despite the lack of previous umbrella review assessing the health outcomes of *Aloe vera*, this type of review has been performed through a wide portion of medical fields. For example, Belbasis et al (18) summarize the environmental risk factors that have been studied about the onset of multiple sclerosis by using an umbrella review approach. They examined 44 unique meta-analyses then reported that the IgG seropositivity to Epstein-Barr virus nuclear antigen (EBNA), infectious mononucleosis, and smoking were the strongest consistent evidence of an association to the onset of multiple sclerosis. Additionally, Dragioti et al (22), assessed the association between antidepressant uses with adverse health

outcomes from 45 meta-analyses of observational studies, they suggested that most of putative adverse health outcomes associated with antidepressant use may not be supported by convincing evidence, and confounding by indication may alter the few associations with convincing evidence.

Not only for synthesizing the association of risk factors with any diseases, but the umbrella review approach can also be used to combine data when many treatments are considered (17). For example, Chakranon et al (23), performed an umbrella review to evaluate the effectiveness of distal technology in multiple health outcomes among people with diabetes which summarized data from 95 articles, including 162 meta-analyses of 46 unique outcomes. Findings from this umbrella review help support the use of distal technology to improve glycemic control, but the evidence is unclear whether distal technology can help improve patient-reported outcomes including quality of life, self-efficacy, and medication-taking. Moreover, conducting an umbrella review can help confirm the need for more robust research to evaluate wider outcomes of distal technology.

To summarize and assess all existing pieces of evidence in the effectiveness of non-pharmacological treatments, herbals, or complementary alternative medicines, the umbrella review approach is also widely used. For example, Wan et al (24), included 16 articles with 50 unique outcomes in their umbrella review to assess the evidence of the various health benefits of allium vegetable consumption, suggested benefit of garlic in cancer prevention and recommended the long-term use as a dietary supplement for patients with dyslipidemia, diabetes, and hypertension. Dinu et al (25) presented an umbrella review of 29 meta-analyses, found the a robust evidences supported association between the adherence to the Mediterranean diet and a reduced the risk of overall mortality, cardiovascular diseases, coronary heart disease, myocardial infarction, overall cancer incidence,

neurodegenerative diseases and diabetes, but they reported that only suggestive or weak evidences supported benefit in most of the site-specific cancer.

5.2 Tools for quality assessment

Bias can arise at all stages of the review process while the reviewer conducted a systematic review and meta-analysis. Thus, it is important to appraise these potential biases when interpreting the results and conclusions from systematic reviews and meta-analyses. In an attempt to assess biases, the Risk of Bias Assessment Tool for Systematic Reviews (ROBIS), and A MeaSurement Tool to Assess systematic Reviews (AMSTAR) has been developed (102, 103). AMSTAR 2 is a revised version of originally published AMSTAR which is one of the most widely used instruments for critically appraising systematic reviews of randomized controlled clinical trials. AMSTAR 2 has been developed to enable a more detailed assessment of systematic reviews that include a randomized controlled trial (RCT) or non-RCTs of health care interventions and also assess the quality of the studies included in a meta-analysis. The AMSTAR 2 has provide a comprehensive user guide for answering 16 questions in its checklist, and has an overall rating based on weaknesses in critical domains (103).

The details of AMSTAR 2 and ROBIS are described in Table 1. ROBIS is an instrument designed for assessing the risk of bias in systematic reviews (rather than in primary studies) (102). ROBIS differs from AMSTAR 2 as it was designed for evaluating the risk of bias specifically, while AMSTAR 2 focuses on a broader goal which is the assessment of the methodological quality of systematic reviews. Moreover, ROBIS could be used in most types of research questions, including diagnosis, prognosis, and etiology but AMSTAR 2 could be used specifically for the reviews of healthcare interventions (103, 104).

Despite the differences in the main domains and questions, some items considered in both AMSTAR 2 and ROBIS were overlapped. Similarly, both AMSTAR 2 and ROBIS can be used to evaluate the risk of bias and methodological quality of systematic reviews of randomized and nonrandomized controlled trials. By using the developed checklist, both 2 tools should not be used to generate overall score. These tools similarly categorize findings into high to low confidence, or unclear in ROBIS.

Table 1 General detail of AMSTAR 2 and ROBIS tools

	AMSTAR 2	**ROBIS**
Aim of development	Assess the methodological quality of systematic reviews	Assess the risk of bias in systematic reviews
Developer	Bruyère Research Institute, Canada	NIHR CLAHRC West, University Hospitals Bristol NHS Foundation Trust, USA
Release Date	2017	2016
Implementation	Systematic reviews including RCT or non-RCTs of health care interventions, or both	Systematic reviews covering questions relating to effectiveness (interventions), etiology, diagnosis, and prognosis.
Answers in the checklists	3 types of answers; yes, no or partial yes (in some items)	5 types of answers; yes, probably yes probably no, no, or no information
Target users	Any reviewers	Any reviewers, some methodologic and/or content expertise would be required

	AMSTAR 2	ROBIS
Characteristics of the tools	16 items (see table 5) with 7 critical domains including; ● Protocol registered before commencement of the review ● Adequacy of the literature search ● Justification for excluding individual studies ● Risk of bias from individual studies being included in the review ● Appropriateness of meta-analytical methods ● Consideration of risk of bias when interpreting the results of the review ● Assessment of presence and likely impact of publication bias	The tool is completed in 3 phases: 1. Assess relevance (optional) 2. Identify concerns with the review process, consists of 4 domains which bias may be introduced into a systematic review including; ● Study eligibility criteria ● Identification and selection of studies; ● Data collection and study appraisal ● Synthesis and findings 3. Assess the overall risk of bias in review and whether this considered limitations identified in any of the phase 2 domains.
Results	*Not* intended to generate an overall score	*Not* be used to generate a summary quality score
Overall rating	4 categories; high, moderate, low, or critically low confidence	3 categories; low, high, or unclear risk of bias

Two studies were conducted to compare ROBIS with the AMSTAR instrument. Buehn et al (105) reported that ROBIS has fair reliability and good construct validity to assess the risk of bias in systematic reviews and Banzi et al (106), found that AMSTAR and ROBIS offer similar inter-rater reliability (IRR). AMSTAR has much better agreement among raters compared to ROBIS. Pieper et al (104), compared AMSTAR 2 with ROBIS in the evaluation of systematic reviews that include both RCTs and non-

RCTs. This study suggested that reliability (reported in IRR) of both tools were fair, with the slightly higher reliability for AMSTAR 2 than for ROBIS (AMSTAR 2: k = 0.30, 95% CI: 0.17 to 0.43; ROBIS: k = 0.28, 95% CI: 0.13 to 0.42). The authors also reported very high correlation between the overall ratings of AMSTAR 2 and ROBIS (Spearman rs = 0.84), suggesting validity. Moreover, the authors stated that raters found AMSTAR 2 easier to apply than ROBIS, with questions more clear, simple and specific. Besides, the AMSTAR 2 guidance was found to be clearer and simpler than the ROBIS guidance; this would probably facilitate its use also by nonexperience reviewers. In terms of usability, the developers of AMSTAR 2 reported that it took them between 15 and 32 minutes (excludes the reading time) to apply AMSTAR 2 (103). In contrast to Pieper et al (104), they reported 18 minutes on average, including reading time, which is much faster.

In conclusion, AMSTAR 2 and ROBIS seemed not to be much different. Both tools can be effectively used to assess the quality of systematic review and meta-analysis. However, AMSTAR 2 is easier for implementation than ROBIS, especially if reviewers are not expertise. Additionally, compared to ROBIS, AMSTAR 2 and the older AMSTAR has been used in the greater numbers of previously published umbrella reviews (21-24, 107, 108).

5.3 Measures of effect size

Effect size, which calculated to summarize the effect of each included study, help indicate the direction and magnitude of the difference between groups (i.e. do the results favor the treatment or control, and if so, by how much) (109). In a meta-analysis, the effect size of each study is calculated. The measurements of effect size that commonly used as follows;

5.3.1 Effect measures for continuous outcome variables:

The measurements of effect size that commonly used for continuous outcome variables are mean difference (MD) and standardized mean difference (SMD). As described in the Cochrane handbook for systematic reviews of intervention, the mean difference, also known as weight mean difference (WMD), is a standard statistic that measures the absolute difference between the mean value in two groups in a clinical trial. It estimates the amount by which the experimental intervention changes the outcome on average compared with the control. It can be used as a summary statistic in meta-analysis when outcome measurements in all studies are made on the same scale. While the standardized mean difference is used as a summary statistic in meta-analysis when the studies all assess the same outcome but measure it in a variety of ways (for example, all studies measure depression but they use different psychometric scales). In this circumstance it is necessary to standardize the results of the studies to a uniform scale before they can be combined (88).

$$\text{SMD} = \frac{Difference\ in\ mean\ outcome\ between\ groups}{Standard\ deviation\ of\ outcome\ among\ participants}$$

5.3.2 Effect measures for binary (dichotomous) outcome variables:

Dichotomous (binary) outcome data arise when the outcome for every participant is one of two possibilities, for example, dead or alive, or clinical improvement or no clinical improvement. This section considers the possible summary statistics when the outcome of interest has such a binary form. The general effect measures that commonly used in clinical trials with dichotomous data are rate or risk ratio (RR), rate or risk difference (RD), and odds ratio (OR) (88). For example, if each study compares a treatment groups with a control groups, the data can be displayed in the form of 2 x 2 table as;

	Event (Success)	No event (Fail)	Total
Experimental (treatment) group	SE	FE	NE
Control group	SC	FC	NC

where SE, SC, FE and FC are the numbers of participants with each outcome ('S' or 'F') in each group ('E' or 'C'). The following summary statistics can be calculated (88):

$$\text{RR} = \frac{risk\ of\ event\ in\ experimental\ group}{risk\ of\ event\ in\ control\ group} = \frac{S_E/N_E}{S_C/N_C}$$

$$\text{OR} = \frac{odds\ of\ event\ in\ experimental\ group}{odds\ of\ event\ in\ control\ group} = \frac{S_E/F_E}{S_C/F_C}$$

RD = risk of event in experimental group – risk of event in control group

$RD = (S_E / N_E) - (S_C / N_C)$

In umbrella review, details of each effect size of included systematic review and meta-analysis should be presented to provide an effective overview. The umbrella review authors can prefer presenting the results of each meta-analysis that included in their review without altering the analysis concept or re-analyze all effect sizes with the same metric (e.g. RR, OR, and HR) (17).

5.4 Statistical models for umbrella review

The statistical models for conducting meta-analysis in umbrella reviews that often used were fixed or random-effect models.

5.4.1. The fixed-effects model

The fixed-effects model assumes that the true effect size is the same in all studies and there is no heterogeneity between the studies. All observed

differences in the data reflects sampling error or chance within each study. According to the statistical stringent of this model, it should be used when the heterogeneity is small determining by Chi-square or I^2 tests.

5.4.2 The random-effects model

The random-effects model allows for more flexibility. The simplest and well-known version of this model is the DerSimonian and Laird method (110). The random-effects model assuming that there may be other factors influencing the data than error or chance, within and between studies. This model assesses both within-study variability and between-study variability. The random-effects model can help provide the generalizations beyond the population included in the studies.

To date, no consensus has been made whether fixed or random-effects models should be used in meta-analysis. In the case that heterogeneity was not present, these two models tend to give similar overall results. However, if heterogeneity was present, the random-effects model could help provide a more conservative estimate of the overall effect size, and the detection of significant differences is less possible.

5.5 Heterogeneity

Heterogeneity refers to variability among studies in a systematic review. Differences in the characteristics of participants, interventions, and outcome measurements may be represented as clinical heterogeneity. Differences in study designs and methodological quality (risk of bias) usually referred as methodological heterogeneity. Besides, the variation of effect sizes between studies may be referred to as statistical heterogeneity. The statistical heterogeneity may occur due to clinical or methodological heterogeneity, or simply by chance (88).

Heterogeneity is a measurement of the relative consistency or inconsistency of studies pooled in a meta-analysis. If a low overlap in the confidence intervals

among primary studies is present, this can indicate heterogeneity (111). A formal statistical test of the between-study heterogeneity is provided by the test of homogeneity. The test can be evaluated using the Cochran's Q test and the I^2 statistic with 95% CI.

The I^2 statistic is the percentage of observed total variation across due to heterogeneity and not due to chance. I^2 ranges between 0% and 100%. According to the Cochrane handbook (88), I^2 exceeding 50% or 75% are indicative of high or very high heterogeneity, respectively (88). Because of this low power, some review authors use a significance level of P-value less than 0.1, rather than the conventional 0.05 value, in order to protect against the possibility of falsely stating that there is no heterogeneity present. When heterogeneity was reported by a meta-analysis as Q or X^2, it can be converted to I^2 with the formula [96]:

$$I^2 = \left(\frac{Q-df}{Q}\right) x\ 100\%$$

Where Q is the chi-squared statistic and df is its degrees of freedom, which equal to k-1.

5.6 Prediction interval

A prediction interval (PI) is a range that presents the expected range of true effects in similar studies. The prediction intervals have been suggested to be routinely reported in addition to the CI to allow more informative inferences in meta-analyses (112). A 95% PI evaluates the expected true effects for 95% of similar (exchangeable) studies that might be conducted in the future. When the PI includes the null (i.e. odds ratio of 1), this suggests that future studies might find results indicating that the exposure produced no effect or the opposite effect on the outcome under consideration (112). However, reporting the PI might have some limitations. For example, the PI may be uncertain if the participants in the future studies are far different from the participants in all studies that have been done in

the past. If the estimates of the summary effect and the heterogeneity are imprecise (e.g. small number of studies included in the meta-analysis), the PI will imprecise too.

If the between-study heterogeneity is not detected, the prediction interval usually concurs with the respective confidence interval (CI). Nevertheless, if the heterogeneity is detected, a PI will cover a wider range than a CI. This occurred in over 70% of the statistically significant meta-analyses in which heterogeneity was detected in the Cochrane Database of Systematic Reviews. Accordingly, in case of a statistically significant effect (where all values of the 95% CI are on the same side of the null), the corresponding 95% PI may indicate that values are possible on both sides of the null. This means that there will be settings where conclusions based on CIs will not hold. In addition, the prediction interval can be used to evaluate the probability that the treatment in a future setting will have a true-positive or true-negative effect.

The PI for the treatment effect in a new trial can be calculated by using the following formula;

$$mean \pm t_{df} \times \sqrt{(se^2 + \tau^2)}$$

where t is the appropriate centile point (e.g., 95%) of the t distribution with k–2 degrees of freedom, se^2 is the squared standard error, and $\boldsymbol{\tau}^2$ the between-study variance.

In the STATA program, the approximate prediction interval will display in the forest plot while calculating the effect sizes, by using the option rfdist. If less than 2 studies were included in the meta-analysis, the distribution is inestimable and effectively infinite. The coverage (e.g., 90%, 95%, or 99%) for the interval may be set by using the command rflevel (#).

5.7 The test for excess significance

The test for excess significance was developed aiming to determine whether there is a relative excess of formally significant findings. The main concept of this test is to investigate in a body of evidence whether the observed (O) number of statistically significant results (positive studies, $p < 0.05$) is larger compared to their expected (E) number of studies with statistically significant results (113). If the observed number of studies with statically significant results in the literature is more than the expected number, it suggests strong biases which possibly caused by selective analyses and selective outcome reporting (114). The evaluation of excess statistical significance can be performed following these steps;

6.3.1. Estimate the effect size: It can be safely assumed that the true effect size is the same in all studies on the same question if the true effect size for any meta-analysis is not known. The effect size can be estimated by using the effect size of the largest studies (e.g. smallest standard error) in each meta-analysis (113).

6.3.2. Calculate the statistical power of each study by using the *power* command in Stata (College Station, TX).

6.3.3. Calculate the expected number of studies with statistically significant results in each meta-analysis by the sum of the statistical power estimates for each component study. The sum of the power estimates gives the expected number of positive studies. The estimated power of each component study depends on the plausible effect size of the largest study in each meta-analysis (115).

$E = \sum_{i=1}^{n}(1 - \beta_i)$ when E = expected number of studies

n = number of published studies

$1 - \beta_i$ = power at the α = 0.05

6.3.4. Compare the expected (E) against observed (O) number of positive studies through the chi-square (χ^2) test; A = $[(O-E)^2 / E + (O-E)^2 / (n-E)]$. The test can be applied regardless of whether the study outcome of interest is binary or continuous.

The power to detect a specific excess may be low, primarily when few 'positive' studies were detected in the meta-analysis. Therefore, excess statistical significance should be noted if two tailed P-value is less than 0.10 as in proposed publication bias tests (113).

5.8 Small-study effects

Small-study effects are the event that smaller studies tend to give different, often larger, intervention effects than larger studies (17). Small-study effects may help reflect publication and other selective reporting biases. Moreover, small-study effects may also exhibit the heterogeneity, chance, or other reasons for differences between small and large studies.

Evidence of small-study effects can be investigated by using the Egger's regression asymmetry test (116). The main idea behind the test for small study-effects is to determine whether there is a statistically significant association between the effect sizes and their measures of precision such as effect-size standard errors. A significance threshold P ≤ 0.10 with more conservative effect in larger studies is considered evidence for small-study effects. For each eligible meta-analysis, Ioannidis and Trikalinos (117) provide the following criteria to ensure whether applying an asymmetry test may be meaningful or appropriate;

- The number of included studies ≥ 10 is considered sufficient, or if not, the tests may have low power;
- There should be at least one study with a statistically significant result ($p < 0.05$);

- There should be no significant heterogeneity (I^2 < 50%). If not, the asymmetry of the funnel plot may be induced by between-study heterogeneity rather than publication bias;
- The difference in precision of the largest and the smallest study was sufficiently large (ratio of the maximum to minimum variances across studies > 4. If it is violated, the funnel plot will look more like a horizontal line than an inverted funnel, and the funnel-asymmetry tests will have an inflated type I error.

The results of the tests of small-study effects should be interpreted with caution. If small-study effects were detected, publication bias together with other reasons should be explored to explain its present. However, even though no small-study effects were detected from the test, this effect may still exist due to the low power of the test method.

6. Evaluation of the certainty of evidences in umbrella review

6.1 Grading of Recommendations Assessment, Development and Evaluation (GRADE)

The GRADE approach is an instrument for rating the quality of evidence. It has been used for appraising quantitative evidence in clinical practice guidelines (118-120), health care recommendations (121), systematic reviews and meta-analyses (122), and umbrella reviews (23, 123). GRADE categorizes the quality of evidence into four levels including high, moderate, low, and very low quality, as shown in Table 2. Theoretically, the authors of the systematic reviews should conduct the GRADE assessment by themselves because they are familiar with the study-level details that are needed for estimating the risk of bias and other details needed for the GRADE assessment (i.e. consistency, and reporting bias in each review case).

Table 2 Four levels of evidence according to GRADE approach

Quality level	Current definition (124)
High	We are very confident that the true effect lies close to that of the estimate of the effect
Moderate	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

6.2 Credibility assessment

According to previously published umbrella reviews, criteria for credibility assessment has been used to classify the level of evidence (18-20, 22, 23, 25, 125). Several umbrella reviews of meta-analyses in various clinical medical areas (i.e. neurology, oncology, nutrition medicine, internal medicine, and psychiatry), stratified the evidence using criteria for credibility assessment. This classification has been performed widely and strongly recommended because they allow an objective, standardized classification of the level of evidence (15). The criteria for credibility assessment have been variably proposed among published umbrella reviews, based upon the following strict criteria using several of statistical data; number of participants included in the meta-analysis, namely p-value (statistically significant of summary effect sizes), prediction interval, heterogeneity, and the presence of biases (e.g. small-study effect, excess-significance bias).

Regarding previously published umbrella reviews, they classified evidence from meta-analyses into 4 categories (convincing, highly suggestive, suggestive, weak) and no-significant associations as follows;

- Convincing (class I) when, included more than 1000 cases, strong statistical significance of $P < 10^{-6}$ or 10^{-3}, had the largest study reporting a significant result ($P < 0.05$), 95% prediction interval excluding the null, did not have large heterogeneity ($I^2 < 50\%$), absence of excess significance bias and small-study effect.
- Highly suggestive (class II) when number of cases > 1000, $p < 10^{-6}$ or 10^{-3}, largest study reporting a significant result ($P < 0.05$) and class I criteria not met.
- Suggestive (class III) when number of cases > 1000, $p < 10^{-3}$ and class I–II criteria not met.
- Weak (class IV) when $p < 0.05$ and class I–III criteria not met;
- Non-significant when $p > 0.05$.

Both GRADE and credibility assessment are tools developed aiming to assess certainty, or strength, of the studies included in the body of evidence, as shown in Table 3. These 2 tools also use some overlapping criteria, such as heterogeneity in credibility assessment, which is considered as inconsistency in GRADE.

Table 3 General details of GRADE and credibility assessment

	GRADE	Credibility assessment
Objective	To assess the certainty, or strength, of the studies included in the body of evidence	To assess the certainty of evidence included in umbrella review
Criteria	1. Evidence from randomized controlled trials (RCT) starts at high quality and evidence that includes observational studies starts at low	1. Number of cases/participants included in meta-analysis 2. Namely p-value (strength of the association)

	GRADE	Credibility assessment
	quality 2. Using 5 factors to decrease and 3 to increase the quality rating as follows; 2.1) Decrease the quality ● Study limitation (risk of bias) ● Inconsistency (i.e. I^2 statistic) ● Indirectness of evidence ● Imprecision ● Publication bias (i.e. visual inspection, asymmetry of funnel plots) 2.2) Increase the quality ● Large magnitude of effect ● All plausible residual confounders or biases would reduce a demonstrated effect ● Dose-response gradient	3. Prediction interval 4. Heterogeneity (e.g. I^2 statistic) 5. Presence of biases (e.g. small-study effect, excess-significance bias)
Classification	4 levels: high, moderate, low, and very low quality	4 categories: convincing, highly suggestive, suggestive, weak and no-significant associations

Although umbrella reviews are performed widely, guidance on how to estimate the certainty of the evidence is relatively limited. While most of the previously published umbrella reviews assessed the certainty of evidence by using the credibility assessment, GRADE has been used infrequently, as shown in Table 4. The reason might be because GRADE is relatively subjective. On the contrary, the criteria for credibility assessment has been proposed specifically for an umbrella review of meta-analyses. Besides, some statistical tests which widely performed in umbrella reviews of the meta-analyses, such as prediction interval and excess

significant bias, are considered as one of the criteria in credibility assessment, but not in GRADE.

In conclusion, although both GRADE and credibility assessment are tools for grading evidence, credibility assessment appears to be more objective because it uses several statistical tests to assess different type of bias and it can work for many types of research questions. It is well known, that heterogeneity, publication bias, small-study effects, and excess of significant bias in the published meta-analyses can contribute to biased results of a meta-analyses. In addition, the criteria for credibility assessment have been performed in a greater number than GRADE in previously published umbrella reviews of meta-analyses. Therefore, as the main analysis in our umbrella review is focusing on meta-analyses, the criteria for credibility assessment seems will be used.

Table 4 Evidence grading tool in previous umbrella reviews

Authors, year published	Type of included studies in umbrella review			Tools for evidence grading	
	observational	interventional		GRADE	credibility assessment
		RCT	non-RCT		
Theodoratou, 2014 (125)	✓	✓			✓
Belbasis, 2015 (18)	✓				✓
Li, 2017 (126)	✓	✓			
Kalliala, 2017 (20)	✓	✓			✓
Poole, 2017 (123)	✓	✓	✓	✓	
Kyrgiou, 2017 (19)	✓				✓
Veronese, 2018 (108)	✓				✓
Dinu, 2018 (25)	✓	✓			✓
Rezende, 2018 (127)	✓				✓
Radua, 2018 (128)	✓				✓
Dragioti, 2018 (21)		✓	✓		✓
Giannakou, 2019 (107)	✓				✓

Authors, year published	Type of included studies in umbrella review			Tools for evidence grading	
	observational	interventional		GRADE	credibility assessment
		RCT	non-RCT		
Yu, 2019 (129)	✓				✓
Dragioti, 2019 (22)	✓				✓
Chakranon, 2019 (23)	✓	✓		✓	✓

CHAPTER III

METHODOLOGY

1. Protocol registration

This umbrella review was done in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines for meta-analyses of interventional studies (90), following a predetermined published protocol (PROSPERO registration: CRD42020152522).

2. Search strategy

The following 6 bibliographical databases were searched from inception until October 2019;

- PubMed
- Excerpta Medica database (Embase)
- Scopus
- The Cochrane database of systematic reviews
- Cumulative Index to Nursing and Allied Health Literature (CINAHL)
- Allied and Complementary Medicine Database (AMED)

The following search terms were used; (‘aloe’ OR ‘aloe vera’) AND (‘systematic review’ OR ‘meta-analysis’). The searches were applied without restrictions on population, study settings, and languages. The full search strategies are described in Appendix 1. References from eligible systematic reviews were also manually reviewed to identify additional studies that may not have been retrieved through search strategies. All identified publications were managed with EndNote version X9 (Clarivate Analytics) and Microsoft Excel (Microsoft Corporation).

3. Study selection

Only systematic reviews or meta-analyses of clinical trials investigated the effects of *Aloe vera* on health outcomes were included in this umbrella review. Individual studies conducted on humans that compared *Aloe vera* with any comparator (i.e. placebo, other treatment options) were included. Health outcomes were defined as any of clinical outcomes (i.e., disease severity and physiological parameters) or patient reported outcomes (i.e., quality of life and satisfaction). Duplicated studies and studies which full-text were not available were excluded. A full description of the inclusion/exclusion criteria using PICOS framework can be found in Table 5.

At the full-text assessment, data from relevant articles were extracted to define the population of interest and the main outcome reported of each systematic review and meta-analysis. When two or more systematic reviews and meta-analyses examining the effects of *Aloe vera* on the same population and health outcome were identified, only the systematic reviews and meta-analysis with the largest number of primary included studies were selected, as described previously (18, 22, 130). If two or more articles including the same number of primary studies, only the meta-analysis which contains the greatest number of patients were included. If two or more articles including the same number and set of patients, only the most recent one was selected. This procedure was adopted to avoid overlapping data as much as possible.

Table 5 Full description of inclusion criteria using PICOS framework

PICOS	Eligibility criteria
Study design	Systematic review and/or meta-analysis of clinical trials
Population	Humans. No restrictions regarding the age of participants, ethnicity or specific group.
Intervention	*Aloe vera* No restriction regarding dose, dosage form, frequency, administration, and

PICOS	Eligibility criteria
	duration of treatment.
Comparator	Any comparators (other treatment options or placebo)
Outcomes	Type of health outcomes as follows: • Clinical outcome: efficacy on indication of Aloe vera such as pain reduction in irritatable bowel syndrome, reduction in fasting plasma glucose in diabetes, time to complete wound healing • Patient reported outcomes such as quality of life, satisfaction
Setting	Any setting including urban, rural, hospital, pharmacy, hospitals

4. Data extraction

The full text of potentially eligible articles was read thoroughly then one reviewer (SS) extracted the data, which were confirmed independently by another reviewer (CR). Discrepancies were resolved by discussion with a third reviewer (TW). The following data were extracted using a standardized form: first author, year of publication, study design, number of included studies, study population, sample size, treatment and control conditions, outcome examined, and main findings. In addition, summary effect sizes of each outcome (i.e., mean difference (MD); standardized mean difference (SMD); risk ratio (RR); odds ratio (OR)) along with p-value and corresponding 95% confidence interval (CI) were extracted. Furthermore, the results of heterogeneity measures (I^2 and its p-value) and Egger's test were also extracted. Lastly, we extracted the effect sizes and their 95% CI of primary studies that were included in individual meta-analyses.

5. Methodological quality assessment

Methodological quality of each included meta-analyses was independently assessed by two reviewers using the AMSTAR 2 (A Measurement Tool to Assess Systematic Reviews) checklist (103), see Appendix 2 for the full checklist.

Any disagreement was resolved by discussion or consultation with a third reviewer. The AMSTAR 2 checklist measures 16 questions, as shown in Table 6. Each of the questions in the checklist can be answered as being 'Yes', 'No', or 'Partial yes'.

Table 6 List of questions from AMSTAR 2 checklist

Item	Question
1	Did the research questions and inclusion criteria for the review include the components of PICO?
2	Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?
3	Did the review authors explain their selection of the study designs for inclusion in the review?
4	Did the review authors use a comprehensive literature search strategy?
5	Did the review authors perform study selection in duplicate?
6	Did the review authors perform data extraction in duplicate?
7	Did the review authors provide a list of excluded studies and justify the exclusions?
8	Did the review authors describe the included studies in adequate detail?
9	Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?
10	Did the review authors report on the sources of funding for the studies included in the review?
11	If meta-analysis was performed, did the review authors use appropriate methods for statistical combination of results?
12	If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?
13	Did the review authors account for RoB in primary studies when interpreting/discussing the results of the review?
14	Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?
15	If they performed quantitative synthesis did the review authors carry out an

Item	Question
	adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?
16	Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

After finishing the appraisal using the AMSTAR 2 checklist, the overall rating was performed based on weaknesses in critical domains. The seven domains of AMSTAR 2 instrument are considered as critical domains that can critically affect the validity of a review and its conclusion as follows (103);

- Protocol registered before commencement (item 2)
- Adequacy of the literature search (item 4)
- Justification for excluding individual studies (item 7)
- Risk of bias from individual studies being included in the review (item 9)
- Appropriateness of meta-analytical methods (item 11)
- Consideration of risk of bias when interpreting the results of the review (item 13)
- Assessment of presence and likely impact of publication bias (item 15)

The methodological quality of each study was rated as high, moderate, low, or critically low using the proposed scheme for interpreting weaknesses detected in critical and non-critical items, as shown in Table 7.

Table 7 AMSTAR 2: Rating overall confidence and interpretation

Rating	Number of weakness	Interpretation
High	0 -1 non-critical weakness	The systematic review provides an accurate and comprehensive summary of the results of the available studies that address the question of interest
Moderate	> 1 non-critical weakness*	The systematic review has more than one weakness but no critical flaws. It may provide an

Rating	Number of weakness	Interpretation
		accurate summary of the results of the available studies that were included in the review
Low	1 critical weakness, with or without non-critical weaknesses	The review has a critical flaw and may not provide an accurate and comprehensive summary of the available studies that address the question of interest
Critically low	> 1 critical weakness, with or without non-critical weaknesses	The review has more than one critical flaw and should not be relied on to provide an accurate and comprehensive summary of the available studies

* Multiple non-critical weaknesses may diminish confidence in the review and it may be appropriate to move the overall appraisal down from moderate to low confidence.

6. Data synthesis and analysis

For systematic reviews, descriptive analyses and present the authors' conclusions was performed. The main analysis of this umbrella review focused on quantitative synthesis from meta-analyses. All analyses were conducted using STATA version 15.0 (College Station, TX) as follows;

6.1 Effect sizes

The effect sizes of each outcome from the included meta-analysis was repooled by using the DerSimonian and Laird random-effects model (110) which calculated by the *metan* command in STATA program. The effect sizes of each outcome included in this umbrella review were reported in the same metric (e.g. RR, OR, and MD) as reported by the original article. The p-value and 95% confidence interval were also calculated along with the effect sizes. A p-value less than 0.05 was considered as statistical significance.

6.2 Heterogeneity

Heterogeneity is a measurement of the relative consistency or inconsistency of studies pooled in a meta-analysis. The heterogeneity was evaluated using the I^2

statistic. I^2 ranges between 0% and 100%. The I^2 ranges of 50% and 75% corresponded to high and very high heterogeneity, respectively.

6.3 Prediction interval (PI)

The 95% PI was calculated, which estimated the interval of future effect size in a new study addressing the same outcome and accounting for between-study variations (131). In STATA program, the 95% prediction interval can be obtained from the forest plot while setting the coverage using the *rflevel(#)* command.

6.4 Evidence of small-study effects

A small-study effect was assessed by the Egger's test to investigate whether small studies tend to give larger effect size than large studies (116). The command used in STATA for egger's test was *metabias*. A significance threshold $P \leq 0.10$ was used.

6.5 Excess significance bias

The existence of excess significance bias were assessed, which evaluates whether the number of observed studies with statistically significant results differs from the expected number of positive studies (113, 114). If there are many studies with statistically significant results in the literature, bias findings might be present from selective analyses and selective outcome reporting. The excess significance test was performed as follows:

1. Estimate the effect size: The plausible effect size in this umbrella review was assumed to be the effect size of the largest studies (e.g. smallest standard error) in each meta-analysis. Due to the true effect size for any meta-analysis is not known, it can be safely assumed that the effect is the same in all studies on the same question (113).

2. The statistical power of each study were calculated by using the *power* command in STATA (College Station, TX).

3. The expected number of studies with statistically significant results in each meta-analysis were calculated by the sum of the statistical power estimated for each component study (115).

$E = \sum_{i=1}^{n}(1-\beta_i)$ E = expected number of studies

n = number of published studies

$1-\beta_i$ = power at the α = 0.05

4. The expected (E) against observed (O) number of positive studies were compared through the chi-square (χ^2) test. These comparisons were done separately for each meta-analysis, then excess statistical significance for individual meta-analyses was considered if P value less than 0.10.

7. Assessment of the Credibility of the Evidence

The credibility assessment criteria were applied as proposed by several previously published umbrella reviews (18, 22, 23, 25). Each of reported health outcome from each meta-analysis was categorized into convincing, highly suggestive, suggestive, weak, or non-significant based on number of cases, statistically significant of summary effect sizes using random effect model (i.e., P ≤ 0.05), prediction intervals excluded the null, presence of large heterogeneity (I^2 > 50%), presence of small-study effects and excess significance bias, as shown in Table 8.

Table 8 Criteria for credibility assessment

Category	Criteria
Convincing (class I)	• Number of cases > 1000 • $p < 10^{-6}$ • No large heterogeneity ($I^2 < 50\%$) • 95% prediction interval excluding the null • No evidence of small-study effects • No evidence of excess significance bias
Highly suggestive (class II)	• Number of cases > 1000 • $p < 10^{-6}$ • Largest study with statistically significant effect • Class I criteria not met
Suggestive (class III)	• Number of cases > 1000 • $p < 10^{-3}$ • Class I-II criteria not met
Weak (class IV)	• $p \leq 0.05$ • Class I-III criteria not met
Non-significant	$p > 0.05$

8. Sensitivity analysis

For each meta-analysis, sensitivity analysis was performed by repeating the primary analysis with an altered dataset by including only RCTs. This analysis aims to determine whether these changes have any effect on the combined outcome estimate and whether the credibility of evidence level changed. In this umbrella review, the sensitivity analyses were performed for evidence that graded as convincing, highly suggestive, or suggestive evidence (class I, II, or III) to determine the robustness of the observed outcomes.

CHAPTER IV

RESULTS

1. Study selection

The literature searching and selection is summarized in figure 2. There were 698 articles identified, of which 582 articles were excluded during screening titles and abstracts and duplicated articles. The potentially relevant full-texts of 116 articles were reviewed. The 96 articles were excluded due to the study design were not systematic reviews and meta-analyses of clinical trials (n = 31), outcomes of *Aloe vera* not reported (n = 12), not the largest studies in each outcome (n = 49), and full-text cannot be retrieved (n = 4). List of excluded studies and reasons for exclusions were described in detail in appendix 3. Finally, 20 articles were eligible including 10 qualitative systematic reviews (132-141) and 10 meta-analyses with 71 unique outcomes (11, 12, 14, 93, 97, 98, 142-145). The included articles published in 2005 to 2019, which all conducted in patients with indications of *Aloe vera*, except 1 study conducted in healthy participants (132).

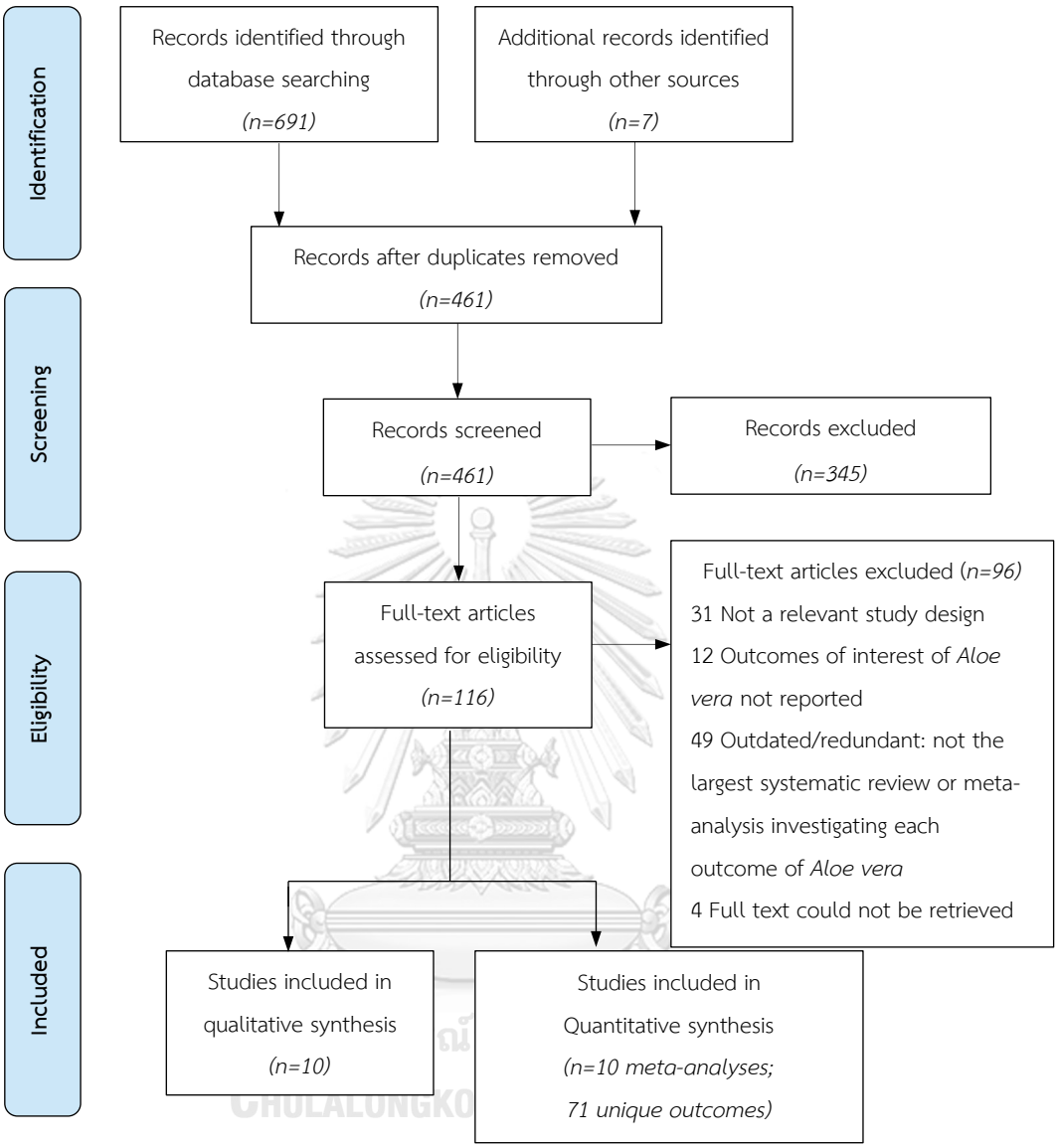

Figure 2 Evidence search and selection

2. Description and Summary of *Aloe vera* on health outcomes

Overall, *Aloe vera* has been investigated in multiple health outcomes including dentistry, anti-diabetes, lipid-lowering, gastrointestinal disorders, phlebitis, radiation-induced reactions, skin conditions, and wound healing.

Ten systematic reviews of clinical trials that described results without quantitative synthesis were included in this umbrella review. The characteristics and descriptive summary are described in appendix 4. In total, 2,039 participants from 26 primary studies were included in these systematic reviews. The median number of primary studies included in these systematic reviews was 1 (IQR 1-4). The median duration of treatment was 4 weeks (IQR 2.7-5.5) with oral (n = 3; 30%) and topical (n = 7; 70%) dosage forms of *Aloe vera*. The included systematic reviews suggested the potential benefits of *Aloe vera* for improving the psoriasis plaques and severity (134), GERD symptoms (136), increasing the frequency of bowel movement and soften stool in chronic constipation (141), treatment of acute radiation proctitis in patients with breast cancer (135), healing cracked nipples (139), reducing plaque in gingivitis (132), and reducing the lesion of acne vulgaris (140), and reducing sign and symptoms of seborrheic dermatitis (138). However, no differences between *Aloe vera* and other treatments were found in the reduction of oral dryness due to radiotherapy (137), quality of life improvement in inflammatory bowel syndrome (136), and symptom improvement in diabetic peripheral neuropathy (133).

Ten meta-analyses of 71 unique outcomes were included in this umbrella review. In total, 94 of primary studies with 14,352 participants were included in the meta-analyses of these 71 unique outcomes. The median number of included primary studies was 3 (IQR 2-5) and median number of participants was 248 (IQR 132.3-393). The median duration of treatment was 8 weeks (IQR 1-8) with oral (n = 13; 18%) and topical (n = 58; 82%) dosage forms of *Aloe vera*. The characteristics and descriptive summary of included meta-analyses are described in appendix 5.

The health outcomes of *Aloe vera* used that reported in all studies can be categorized into 7 categories including wound healing, radiation-induced mucositis, phlebitis, irritable bowel syndrome, lipid and glucose lowering, and dentistry, as presented in Table 9. Among these 71 outcomes from all included meta-analyses, phlebitis is the most reported outcome (53.5%), as shown in figure 3.

Table 9 Characteristics and quantitative synthesis of the meta-analyses

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Dentistry										
Oral lichen planus (98)										
Pain and burning sensation	3	topical AV/ placebo or corticosteroid	121 (64/57)	WMD	-0.39 (-1.30 to 0.53)	0.41	44.1 (0.17)	-9.85 to 9.08	0.06 (0.01 to 0.11)	non-significant
Clinical improvement	3	topical AV/ placebo or corticosteroid	121 (64/57)	WMD	-0.05 (-0.39 to 0.30)	0.79	37.4 (0.20)	-3.53 to 3.44	0.08 (0.03 to 0.13)	non-significant
Oral submucous fibrosis (97)										
Burning sensation at 1 month	2	topical AV/ placebo or medical interventions	111 (56/55)	WMD	-1.22 (-2.35 to -0.08)	0.04	82.7 (0.02)	NA	-1.75 (-2.24 to-1.26)	weak
Burning sensation at 2 months	2	topical AV/ placebo or medical interventions	111 (56/55)	WMD	-1.33 (-1.95 to -0.72)	2.17×10^{-5}	38.9 (0.20)	NA	-1.62 (-2.25 to-0.99)	weak

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Burning sensation at 3 months	3	topical AV/ placebo or medical interventions	131 (66/65)	WMD	-0.94 (-2.02 to 0.15)	0.09	81.8 (<0.001)	-12.74 to 10.87	-0.37 (-0.75 to 0.01)	non-significant
Mouth opening at 1 month	5	topical AV/ placebo or medical interventions	319 (160/159)	WMD	-0.33 (-2.09 to 1.43)	0.71	89.3 (<0.001)	-6.73 to 6.07	-1.60 (-2.32 to-0.88)	non-significant
Mouth opening at 2 months	5	topical AV/ placebo or medical interventions	393 (197/196)	WMD	-1.23 (-4.28 to 1.83)	0.43	96.7 (<0.001)	-13.05 to 10.60	-2.00 (-2.81 to-1.19)	non-significant
Mouth opening at 3 months	6	topical AV/ placebo or medical interventions	413 (207/206)	WMD	-0.96 (-3.82 to 1.91)	0.51	96.0 (<0.001)	-11.14 to 9.22	-2.00 (-2.81 to-1.19)	non-significant
Tongue protrusion at 1 month	4	topical AV/ placebo or medical interventions	351 (176/175)	WMD	-0.16 (-2.98 to 2.65)	0.91	95.3 (<0.001)	-13.14 to 12.81	-0.40 (-1.06 to 0.26)	non-significant
Tongue protrusion at 2 months	4	topical AV/ placebo or medical interventions	351 (176/175)	WMD	-2.00 (-6.41 to 2.41)	0.37	97.8 (<0.001)	-23.01 to 19.01	-3.00 (-3.74 to-2.26)	non-significant

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Tongue protrusion at 3 months	5	topical AV/ placebo or medical interventions	371 (186/185)	WMD	0.25 (-2.88 to 3.37)	0.88	95.5 (<0.001)	-10.87 to 11.36	-0.40 (-1.06 to 0.26)	non-significant
Cheek flexibility at 1 month	2	topical AV/ placebo or medical interventions	111 (56/55)	WMD	-0.05 (-0.17 to 0.07)	0.40	95.1 (<0.001)	NA	-0.11 (-0.14 to-0.08)	non-significant
Cheek flexibility at 2 months	2	topical AV/ placebo or medical interventions	111 (56/55)	WMD	-0.04 (-0.15 to 0.07)	0.48	88.1 (<0.001)	NA	-0.09 (-0.13 to-0.05)	non-significant
Cheek flexibility at 3 months	3	topical AV/ placebo or medical interventions	131 (66/65)	WMD	-0.02 (-0.13 to 0.09)	0.72	90.8 (<0.001)	-1.33 to 1.29	-0.10 (-0.14 to-0.06)	non-significant
Anti-diabetes										
Glucose lowering in prediabetic and early non-treated DM (12)										
FBG	5	oral AV/placebo	328 (163/165)	WMD	-30.05 (-54.87 to -5.23)	0.02	99.9 (<0.001)	-128.42 to 68.32	-2.40 (-3.01 to-1.79)	weak
HbA1C	2	oral AV/placebo	76 (38/38)	WMD	-0.41 (-0.55 to -0.27)	6.48×10^{-9}	0.0 (0.61)	NA	-0.42 (-0.57 to-0.27)	weak

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Insulin level	2	oral AV/placebo	151 (74/77)	SMD	-1.73 (-4.09 to 0.63)	0.15	96.4 (<0.001)	NA	-1.90 (-2.13 to-1.67)	non-significant
Glucose lowering in type 2 DM (93)										
FBG	5	oral AV/placebo or no treatment	235 (117/118)	WMD	-1.17 (-2.35 to 0.001)	0.05	79.1 (<0.001)	-5.41 to 3.06	-1.30 (-2.37 to-0.23)	weak
HbA1C	4	oral AV/placebo or no treatment	164 (65/99)	WMD	-10.99 (-19.43 to -2.55)	0.01	69.9 (0.02)	-46.99 to 25.01	-9.00 (-16.83 to-1.17)	weak
Dyslipidemia (12)										
TG	4	oral AV/placebo	206 (103/103)	WMD	-43.92 (-66.33 to -21.51)	1.22×10^{-4}	99.8 (<0.001)	-146.21 to 58.37	-11.80 (-12.52 to-11.08)	weak
TC	4	oral AV/placebo	206 (103/103)	WMD	-16.94 (-23.39 to -10.50)	2.53×10^{-7}	91.5 (<0.001)	-45.15 to 11.26	-11.60 (-12.84 to-10.36)	weak
HDL	3	oral AV/placebo	136 (68/68)	WMD	2.67 (0.11 to 5.23)	0.04	85.9 (<0.001)	-25.94 to 31.28	1.70 (1.24 to 2.16)	weak
LDL	3	oral AV/placebo	136 (68/68)	WMD	-13.30 (-17.19 to -9.41)	2.02×10^{-11}	96.2 (<0.001)	-57.39 to 30.79	-15.10 (-15.79 to-14.41)	weak

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Gastrointestinal disorders										
Irritable bowel syndrome (11)										
Symptoms scores improvement	3	oral AV (juice, extract)/placebo	137 (70/67)	SMD	0.42 (0.08 to 0.76)	0.02	0.0 (0.90)	-1.78 to 2.62	1.01 (-0.57 to 2.59)	weak
Short term symptoms scores improvement (at 1 month)	2	oral AV (juice, extract)/placebo	112 (58/54)	SMD	0.40 (0.03 to 0.77)	0.04	0.0 (0.69)	NA	0.47 (-0.03 to 0.97)	weak
Long-term symptoms scores improvement (at 3 months)	2	oral AV (juice, extract)/placebo	67 (36/31)	SMD	0.19 (-0.40 to 0.79)	0.52	30.8 (0.23)	NA	-0.06 (-0.67 to 0.55)	non-significant
Response rate	2	oral AV (juice, extract)/placebo	112 (58/54)	RR	1.60 (1.00 to 2.54)	0.05	0.0 (0.96)	NA	1.59 (0.90 to 2.79)	
Phlebitis										
Chemotherapy-induced phlebitis prevention (142)										
Overall incidence	10	topical AV/ conventional treatment	3983 (2493/ 1490)	OR	0.13 (0.08 to 0.20)	9.68×10^{-20}	50.8 (0.03)	0.04 to 0.43	0.08 (0.06 to 0.11)	highly suggestive

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Incidence of 1st-degree CIP	8	topical AV/ conventional treatment	3925 (2469/ 1456)	OR	0.53 (0.21 to 1.33)	0.18	73.9 (<0.001)	0.03 to 9.46	1.77 (0.93 to 3.37)	non-significant
Incidence of 2nd-degree CIP	8	topical AV/ conventional treatment	3925 (2469/ 1456)	OR	0.10 (0.07 to 0.14)	3.41×10^{-35}	0.0 (0.83)	0.06 to 0.15	0.08 (0.05 to 0.13)	highly suggestive
Incidence of 3rd-degree CIP	8	topical AV/ conventional treatment	3925 (2469/ 1456)	OR	0.10 (0.03 to 0.34)	1.90×10^{-4}	38.5 (0.12)	0.01 to 2.04	0.11 (0.01 to 0.92)	suggestive
Chemotherapy-induced phlebitis treatment (142)										
Overall efficacy rate	6	topical AV/ topical 50% $MgSO_4$	547 (282/265)	RR	1.28 (1.19 to 1.38)	8.10×10^{-11}	0.0 (0.85)	1.15 to 1.42	1.18 (1.02 to 1.37)	weak
Overall cure rate	4	topical AV/ topical 50% $MgSO_4$	293 (152/141)	RR	2.38 (1.27 to 4.47)	0.01	75.2 (0.01)	0.16 to 36.39	1.67 (1.06 to 2.63)	weak
Infusion phlebitis prevention (14)										
Total incidence (treatment duration 5 days)	5	topical AV/ no treatment	532 (266/266)	RR	0.27 (0.07 to 1.09)	0.07	96.8 (<0.001)	0.00 to 58.43	0.76 (0.60 to 0.97)	non-significant

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Total incidence (treatment duration 1-7 days)	2	topical AV/ no treatment	370 (185/185)	RR	0.31 (0.18 to 0.52)	9.52×10^{-6}	37.9 (0.21)	NA	0.37 (0.60 to 0.97)	weak
Total incidence (treatment duration 3 days)	3	topical AV/ no treatment	312 (156/156)	RR	0.43 (0.28 to 0.67)	1.67×10^{-4}	59.7 (0.08)	0.00 to 46.59	0.51 (0.35 to 0.75)	weak
Total incidence (treatment duration 2-3 days)	2	topical AV/ no treatment	189 (95/94)	RR	0.21 (0.05 to 0.83)	0.03	75.6 (0.04)	NA	0.11 (0.05 to 0.26)	weak
Incidence of 2nd-degree phlebitis	14	topical AV/ no treatment	4585 (2791/ 1794)	RR	0.18 (0.10 to 0.32)	1.75×10^{-9}	71.2 (<0.001)	0.03 to 1.14	0.06 (0.04 to 0.09)	highly suggestive
Incidence of 2nd-degree phlebitis (treatment duration 5 days)	4	topical AV/ no treatment	482 (241/241)	RR	0.19 (0.07 to 0.55)	0.002	0.0 (0.93)	0.02 to 1.95	0.22 (0.05 to 1.00)	weak
Incidence of 2nd-degree phlebitis	2	topical AV/ no treatment	450 (225/225)	RR	0.21 (0.11 to 0.41)	3.24×10^{-6}	0.0 (0.35)	NA	0.25 (0.12 to 0.52)	weak

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
(treatment duration 1-7 days)										
Incidence of 2nd-degree phlebitis (treatment duration 3 days)	3	topical AV/ no treatment	314 (156/158)	RR	0.42 (0.22 to 0.81)	0.01	29.4 (0.24)	NA	0.47 (0.23 to 0.97)	weak
Incidence of 2nd-degree phlebitis (treatment duration 2-3 days)	2	topical AV/ no treatment	189 (95/94)	RR	0.07 (0.01 to 0.57)	0.01	54.5 (0.14)	NA	0.03 (0.00 to 0.18)	weak
Incidence of 3rd-degree phlebitis	14	topical AV/ no treatment	4585 (2791/ 1794)	RR	0.13 (0.05 to 0.34)	2.34×10^{-5}	43.8 (0.05)	0.01 to 1.78	0.29 (0.06 to 1.30)	suggestive
Incidence of 3rd-degree phlebitis (treatment duration 5 days)	4	topical AV/ no treatment	482 (241/241)	RR	0.26 (0.03 to 2.26)	0.22	0.0 (0.82)	NA	0.20 (0.01 to 4.10)	non-significant

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Incidence of 3rd-degree phlebitis (treatment duration 1-7 days)	2	topical AV/ no treatment	370 (185/185)	RR	0.11 (022 to 0.56)	0.008	0.0 (0.79)	NA	0.13 (0.02 to 0.98)	weak
Incidence of 3rd-degree phlebitis (treatment duration 3 days)	3	topical AV/ no treatment	314 (156/158)	RR	0.29 (0.11 to 0.80)	0.02	0.0 (0.99)	0.0 to 209	0.29 (0.06 to 1.30)	weak
Total incidence of phlebitis	2	topical AV/ potato slice	276 (143/133)	RR	0.91 (0.58 to 1.40)	0.65	0.0 (0.76)	NA	0.96 (0.54 to 1.69)	non-significant
Total incidence of 2nd-degree phlebitis	2	topical AV/ potato slice	276 (143/133)	RR	1.14 (0.49 to 2.67)	0.76	0.0 (0.84)	NA	1.21 (0.44 to 3.37)	non-significant
Total incidence of phlebitis	2	topical AV/ topical 33% $MgSO_4$	200 (100/100)	RR	0.43 (0.24 to 0.77)	0.005	0.0 (0.72)	NA	0.40 (0.19 to 0.82)	weak
Total incidence of phlebitis	2	topical AV/ topical 50% $MgSO_4$	248 (136/112)	RR	0.41 (0.16 to 1.07)	0.07	75.5 (0.04)	NA	0.60 (0.45 to 0.79)	non-significant
Total incidence of 2nd -degree phlebitis	2	topical AV/ topical 50% $MgSO_4$	248 (136/112)	RR	0.28 (0.15 to 0.54)	1.29×10^{-4}	0.0 (0.77)	NA	0.30 (0.14 to 0.62)	weak

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Total incidence of 3rd -degree phlebitis	2	topical AV/ topical 50% $MgSO_4$	248 (136/112)	RR	0.27 (0.07 to 0.98)	0.05	0.0 (0.81)	NA	0.24 (0.04 to 1.24)	weak
Infusion phlebitis treatment (14)										
Rate of resolution: marked improvement	2	topical AV/ topical 33% $MgSO_4$	302 (160/142)	RR	1.93 (1.30 to 2.86)	0.001	33.3 (0.22)	NA	2.27 (1.54 to 3.33)	weak
Rate of resolution: total improvement	3	topical AV/ topical 33% $MgSO_4$	422 (220/202)	RR	1.17 (1.08 to 1.28)	2.53×10^{-4}	37.6 (0.20)	0.52 to 2.63	1.13 (1.04 to 1.24)	weak
Rate of resolution: recovery	7	topical AV/ topical 50% $MgSO_4$	595 (310/285)	RR	1.51 (1.24 to 1.85)	4.94×10^{-5}	49.0 (0.07)	0.89 to 2.58	1.14 (0.90 to 1.44)	weak
Rate of resolution: recovery (treatment duration 3 days)	4	topical AV/ topical 50% $MgSO_4$	394 (206/188)	RR	1.48 (1.13 to 1.93)	4×10^{-4}	57.1 (0.07)	0.52 to 4.19	1.14 (0.90 to 1.44)	weak
Rate of resolution: recovery (treatment duration 15 days)	4	topical AV/ topical 50% $MgSO_4$	151 (79/72)	RR	1.43 (1.13 to 1.80)	0.003	0.0 (0.43)	NA	1.35 (1.03 to 1.77)	weak
Rate of resolution: marked improvement	9	topical AV/ topical 50% $MgSO_4$	814 (417/397)	RR	1.61 (1.29 to 2.02)	3.21×10^{-5}	82.5 (<0.001)	0.77 to 3.36	1.29 (1.10 to 1.51)	weak

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Rate of resolution: marked improvement (treatment duration 3 days)	7	topical AV/ topical 50% $MgSO_4$	679 (349/330)	RR	1.64 (1.23 to 2.19)	0.001	86.1 (<0.001)	0.64 to 4.18	1.29 (1.10 to 1.51)	weak
Rate of resolution: total improvement	10	topical AV/ topical 50% $MgSO_4$	880 (453/427)	RR	1.22 (1.16 to 1.29)	$7.11x10^{-14}$	5.9 (0.39)	1.13 to 1.32	1.14 (1.03 to 1.27)	weak
Rate of resolution: total improvement (treatment duration 3 days)	8	topical AV/ topical 50% $MgSO_4$	679 (349/330)	RR	1.20 (1.13 to 1.26)	$1.08x10^{-10}$	0.0 (0.58)	1.11 to 1.29	1.14 (1.03 to 1.27)	weak
Rate of resolution: total improvement (treatment duration 15 days)	2	topical AV/ topical 50% $MgSO_4$	151 (79/72)	RR	1.35 (1.16 to 1.56)	$7.63x10^{-5}$	0.0 (0.61)	NA	1.39 (1.15 to 1.69)	weak
Rate of resolution: recovery	3	topical AV + non-AV medication/ same non-AV	283 (142/141)	RR	1.76 (1.23 to 2.52)	0.002	61.1 (0.08)	0.04 to 87.79	1.32 (0.96 to 1.80)	weak

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Rate of resolution: marked improvement	3	topical AV + non-AV medication/ same non-AV	163 (82/81)	RR	1.26 (1.09 to 1.47)	0.003	0.0 (0.64)	0.47 to 3.39	1.21 (1.00 to 1.46)	weak
Rate of resolution: total improvement	4	topical AV + non-AV medication/ same non-AV	323 (162/161)	RR	1.23 (1.09 to 1.39)	0.001	42.4 (0.16)	0.80 to 1.89	1.33 (1.15 to 1.53)	weak
Radiation-induced reaction (145)										
Mucositis prevention	2	topical AV/placebo	119 (58/61)	RR	0.75 (0.50 to 1.12)	0.16	58.6 (0.12)	NA	0.61 (0.43 to 0.88)	non-significant
Wound healing										
Burns (143)										
Time to wound healing	3	topical AV/SSD or framycetin cream	210 (105/105)	WMD	-7.79 (-17.87 to 2.29)	0.13	94.3 (<0.001)	NA	-2.90 (-4.10 to-1.70)	non-significant
Infection	3	topical AV/ SSD cream	221 (111/110)	RR	0.93 (0.26 to 3.34)	0.92	0.0 (0.44)	NA	0.75 (0.19 to 3.01)	non-significant

Outcome examined	No.of included studies	Type of AV/control	Total n (AV/ control)	Effect metrics	Effect size[a] (95% CI)	*P* value	I^2 (*P* value)	95% PI	Largest study, effect size (95% CI)	Credibility of evidence
Acute surgical wound (144)										
Wound healing number	2	topical AV (gel, dressing, juice)/ conventional treatment	97 (50/47)	RR	16.33 (3.46 to 77.15)	0.0004	3.53 (<0.001)	NA	NR	weak
Mean time to wound healing	2	topical AV (gel, dressing, juice)/ conventional treatment	101 (50/51)	MD	11.03 (-22.17 to 44.24)	0.51	89 (0.003)	NA	NR	non-significant
Chronic wound (144)										
Wound healing number	5	topical AV (gel, dressing, juice)/ conventional treatment	233 (116/117)	RR	1.73 (1.21 to 2.49)	0.003	51 (0.09)	NA	NR	weak

[a] Effect size base on random-effects model.

Abbreviations:, AV – *Aloe vera*, n – number of participants, ES – effect size, I^2 – heterogeneity, CI – confidence interval, PI – prediction interval, WMD – weight mean difference, SMD – standardized mean difference, OR – odds ratio, RR – relative risk, C – control, NA – not applicable, NR – not reported, CIP – chemotherapy induced phlebitis, $MgSO_4$ – magnesium sulfate, SSD – silver sulfadiazine

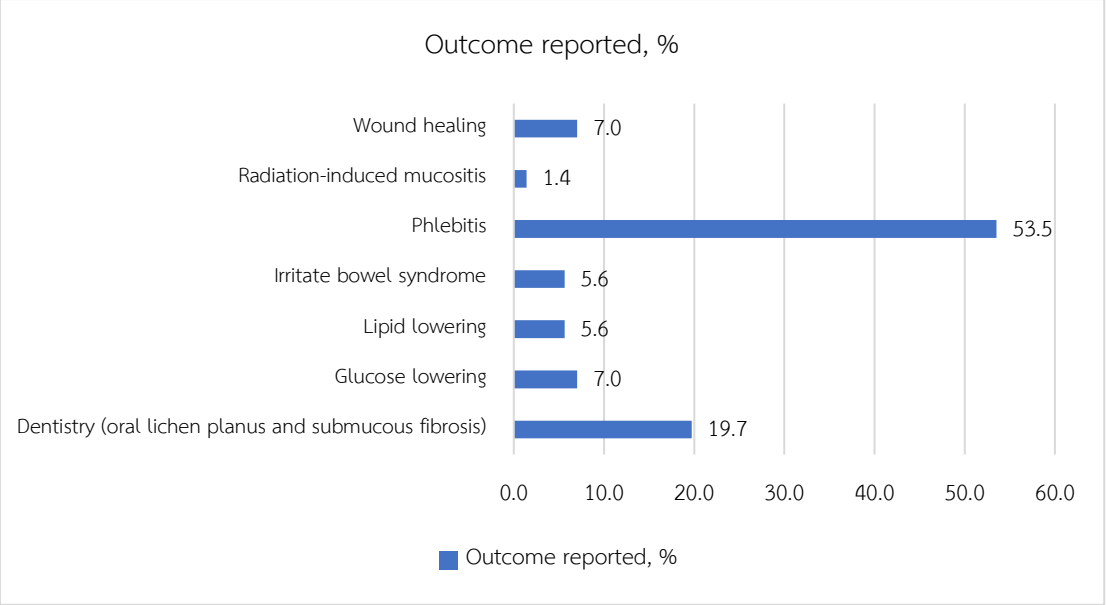

Figure 3 Percentages of the reported 71 outcomes from all studies, classified into 7 categories

Among the 71 outcomes examined, 6 outcomes (8.5%) included more than 1,000 participants (e.g. the prevention of overall, 1st-degree, 2nd-degree, and 3rd-degree CIP, the prevention of 2nd and 3rd-degree infusion phlebitis). The effect sizes of the largest study of 48 outcomes (68%) were statistically significant ($p \leq 0.05$). The 95% predictive intervals excluded the null value for only 5 outcomes (7%), as shown in Table 9 and 10. Data of all reported outcomes were able to reanalyze, except for 3 outcomes from 1 study (144) because detailed data were not available. Additional details of each study included in each meta-analysis are presented in the appendix 6 and 7.

Based on random-effects models, 47 outcomes (67%) were nominally statistically significant at $p \leq 0.05$, 23 (32%) outcomes were significant at $p < 0.001$, and 9 (13%) outcomes reached at p of $< 10^{-6}$. All of the statistically significant outcomes ($p \leq 0.05$) suggested the potential benefits of *Aloe vera.* Among them, most of the statistically significant outcomes suggested the benefit of *Aloe vera* in infusion phlebitis (n=28; 60%). Across all reported, *Aloe vera* showed less effective

than comparator group in only 2 outcomes including the improvement of tongue protrusion in oral submucous fibrosis after 3 months of *Aloe vera* treatment when compared to an antioxidant-control group, and the reduction of second-degree CIP when compared with a potato slice. However, the differences were not statistically significant.

3. Methodological quality assessment

Using the AMSTAR 2 tool, results from 3 meta-analyses were rated as having a high-quality level, 1 as moderate, 2 as low, and 4 as critically low, as reported in Table 10 and the breakdown of answers in each question reported in Table 11.

The majority of the meta-analyses did not meet the AMSTAR 2 critical domains relating to the protocol registration before commencement of the review (n = 6; 60%) and justification for excluding individual studies (n = 4; 40%). The non-critical domains which most of the meta-analyses did not meet requirements (n = 5; 50%) were the domain that relating to identifying the sources of funding for included studies, as shown in figure 4.

4. Small study effects and heterogeneity

Thirty-four (48%) and 26 (37%) outcomes had high (I^2 > 50%) and very high heterogeneity (I^2 > 75%). The Egger's test was performed in 39 (55%) meta-analyses, as the remaining reviews could not be estimated due to insufficient numbers (< 3 primary studies in meta-analysis), indicating small-study effects in 10 (14%) meta-analyses (Egger's test P ≤ 0.10), as shown in Table 10. This included the meta-analyses that examined effect of *Aloe vera* in 6 outcomes of phlebitis treatment (incidence of total improvement, marked improvement, and recovery rate when compared *Aloe vera* with 50% $MgSO_4$), 2 outcomes of chemotherapy-induced phlebitis prevention (incidence of overall and the second-degree CIP), alleviation of

pain and burning sensation in OLP, and mouth opening improvement at 1 month in OSF.

5. Excess significance bias

We further assessed the presence of excess significance bias to determine if the observed number of studies with nominally significant results was different from the expected number ($p \leq 0.10$). Of 71 outcomes, excess significance bias was found in 3 (4%) outcomes, which examined the recovery rate of infusion phlebitis by comparing *Aloe vera* with 50% $MgSO_4$, as shown in Table 10. However, excess significance bias in 34 (48%) of meta-analyses were not present, thus, excess significance bias should be less likely.

6. Credibility of the evidence

The credibility assessment of the 71 outcomes are presented in Table 9 and 10. Among them, only 3 (4%) outcomes were supported by highly suggestive evidence (class II), one of these supported benefits of *Aloe vera* in the prevention of second-degree infusion phlebitis when compared with no treatment with the high methodological quality based on AMSTAR 2 assessment. Two of highly suggestive evidences demonstrated beneficial effects of *Aloe vera* in chemotherapy-induced phlebitis prevention (improvement of overall incidence of CIP and the incidence of second-degree CIP); however, the methodological quality of these meta-analyses reached only the critically low level based on AMSTAR 2.

Suggestive evidence (class III) was found supporting the 2 (3%) outcomes, 1 demonstrated benefit of *Aloe vera* in the prevention of third-degree infusion phlebitis when compared with no treatment, and another supported benefit in prevention of the third-degree CIP.

Among 71 outcomes, majority of the evidence (n = 42; 59%) was weak reporting nominally statistically significant (p-value ≤ 0.05) using a random-effects model. All of these multiple health outcomes supported benefit of *Aloe vera*. For the remaining 24 (34%) outcomes, non-significant evidence (p > 0.05) was found.

6.1 Dentistry

In total, the effects of *Aloe vera* in dentistry were reported in 14 outcomes. Two outcomes of *Aloe vera* in oral lichen planus (OLP) and 12 outcomes in oral submucous fibrosis (OSF) were examined. Most meta-analyses (n = 13) had reported no statistical difference between *Aloe vera* and the control group. Except for 1 meta-analysis suggested the benefit of using topical *Aloe vera* for 2 months in the reduction of a burning sensation among patients with OSF (WMD -1.33; [CI -1.95 to -0.72]; p = 2.17×10^{-5}; I^2 = 38.9; weak credibility of evidence).

6.2 Anti-diabetes

The glucose-lowering effect of *Aloe vera* was reported in 5 outcomes. Three outcomes were investigated in prediabetic and early non-treated diabetic patients while 2 outcomes were investigated in type 2 diabetic patients. *Aloe vera* showed benefit in all group of patients for FBG and HBA_1C reduction. Of these, 1 outcome had high heterogeneity (I^2 > 50%), whereas 2 outcomes had very high heterogeneity (I^2 > 75%). The certainty of evidence was weak.

The lipid-lowering effect of *Aloe vera* in type 2 diabetic patients was reported in 4 outcomes. All of them suggested the benefit of *Aloe vera* which reduced TG, TC, and LDL level, and increased the HDL level. However, the credibility of evidence was weak and all outcomes had very high heterogeneity (I^2 > 75%).

6.3 Gastrointestinal disorders

Effects of *Aloe vera* in gastrointestinal disorders were reported in 4 outcomes that investigated in irritable bowel syndrome (IBS). Of these, 3 outcomes suggested

the statistically significant benefits of *Aloe vera* in the improvement of IBS symptom score in overall duration used and when used for 1 month, and also suggested the benefits of *Aloe vera* in the improvement of response rate. However, the credibility of evidence of these 3 outcomes was weak.

6.4 Phlebitis

In total, the effects of *Aloe vera* in phlebitis were reported in 38 outcomes including infusion phlebitis (n = 32) and chemotherapy-induced phlebitis (n = 6). Most of the meta-analyses (n = 36) had reported the potential benefits of *Aloe vera* in phlebitis prevention and treatment. Only 2 meta-analyses had reported that potato slice (control group) showed higher efficacy than *Aloe vera* in infusion phlebitis prevention, but no statistically significant difference was found. Three outcomes of *Aloe vera* in the phlebitis prevention reached the highly-suggestive level of credibility.

6.5 Radiation-induced reactions

One outcome reported the effect of *Aloe vera* in radiation-induced mucositis. *Aloe vera* showed the potential benefit than placebo but no statistically significance was found (RR 0.75; [CI 0.50 to 1.12]; p = 0.16; I^2 = 58.6; weak certainty).

6.6 Wound healing

In total, the effects of *Aloe vera* in wound healing were reported in 5 outcomes including burn wounds (n = 2), acute surgical wound (n = 2), and chronic wound (n = 1). *Aloe vera* was reported higher efficacy in healing acute-surgical and chronic wounds with a weak level of credibility of evidence. However, *Aloe vera* showed no statistically significant difference in burn wounds healing and infection.

Table 10 Summary of the credibility of evidence and AMSTAR 2 level of meta-analyses reporting the effect of *Aloe vera* on health outcomes (n=71)

				Criteria for credibility assessment							
Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
Outcomes supported by highly suggestive evidence (class II)											
Infusion phlebitis prevention	incidence of 2nd-degree phlebitis (VS no treatment)	Zheng,2014[14] (14)	AV	4585	$<10^{-6}$	Y	N (0.56)	NP	N	Y	High
CIP prevention	incidence of 2nd-degree CIP	Gao,2016[142] (8)	AV	3925	$<10^{-6}$	N	Y (0.04)	NP	Y	Y	Critically low
CIP prevention	overall incidence	Gao,2016[142] (10)	AV	3983	$<10^{-6}$	Y	Y (0.10)	NP	Y	Y	Critically low
Outcomes supported by suggestive evidence (class III)											
Infusion phlebitis prevention	incidence of 3rd-degree phlebitis (VS no treatment)	Zheng,2014[14] (14)	AV	4585	$<10^{-3}$	N	N (0.28)	NP	N	N	High
CIP prevention	incidence of 3rd-degree CIP	Gao,2016[142] (8)	AV	3925	$<10^{-3}$	N	N (0.22)	NP	N	Y	Critically low

Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Criteria for credibility assessment: Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
Outcomes supported by weak evidence (class IV)											
Infusion phlebitis prevention	total incidence (treatment duration 1-7 days) VS no treatment	Zheng,2014[(14)] (2)	AV	370	$<10^{-3}$	N	NA	N (0.85)	NA	Y	High
Infusion phlebitis prevention	total incidence (treatment duration 3 days) VS no treatment	Zheng,2014[(14)] (3)	AV	312	$<10^{-3}$	Y	NA	N (0.84)	N	Y	High
Infusion phlebitis prevention	total incidence (treatment duration 2-3 days) VS no treatment	Zheng,2014[(14)] (2)	AV	189	≤0.05	Y	NA	NP	NA	Y	High
Infusion phlebitis prevention	incidence of 2nd-degree phlebitis (treatment duration 5 days) VS no treatment	Zheng,2014[(14)] (4)	AV	482	≤0.05	N	N (0.96)	NP	N	N	High
Infusion phlebitis prevention	incidence of 2nd-degree phlebitis (treatment duration 1-7 days)	Zheng,2014[(14)] (2)	AV	450	$<10^{-3}$	N	NA	N (0.69)	NA	Y	High

Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Criteria for credibility assessment: Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
	VS no treatment										
Infusion phlebitis prevention	incidence of 2nd-degree phlebitis (treatment duration 3 days) VS no treatment	Zheng,2014[14] (3)	AV	314	≤0.05	N	N (0.52)	N (0.72)	NA	Y	High
Infusion phlebitis prevention	incidence of 2nd-degree phlebitis (treatment duration 2-3 days) VS no treatment	Zheng,2014[14] (2)	AV	189	≤0.05	Y	NA	NP	NA	Y	High
Infusion phlebitis prevention	incidence of 3rd-degree phlebitis (treatment duration 1-7 days) VS no treatment	Zheng,2014[14] (2)	AV	370	≤0.05	N	NA	NP	NA	Y	High
Infusion phlebitis prevention	incidence of 3rd-degree phlebitis (treatment duration 3 days) VS no treatment	Zheng,2014[14] (3)	AV	314	≤0.05	N	NA	NP	N	N	High

Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Criteria for credibility assessment: Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
Infusion phlebitis prevention	total incidence of phlebitis VS 33% $MgSO_4$	Zheng,2014[(14)] (2)	AV	200	≤0.05	N	NA	NP	NA	Y	High
Infusion phlebitis prevention	total incidence of 2^{nd} -degree phlebitis VS 50% $MgSO_4$	Zheng, 2014(14)	AV	248	$<10^{-3}$	N	NA	N (0.42)	NA	Y	High
Infusion phlebitis prevention	total incidence of 3^{rd} -degree phlebitis VS 50% $MgSO_4$	Zheng,2014[(14)] (2)	AV	248	≤0.05	N	NA	NP	NA	N	High
Infusion phlebitis treatment	rate of resolution: marked improvement VS 33% $MgSO_4$	Zheng,2014[(14)] (2)	AV	302	≤0.05	N	NA	NP	NA	Y	High
Infusion phlebitis treatment	rate of resolution: total improvement VS 33% $MgSO_4$	Zheng,2014[(14)] (3)	AV	422	$<10^{-3}$	N	N (0.31)	N (0.29)	N	Y	High
Infusion phlebitis treatment	rate of resolution: recovery VS 50% $MgSO_4$	Zheng,2014[(14)] (7)	AV	595	$<10^{-3}$	N	Y (<0.001)	Y (<0.001)	N	N	High

Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Criteria for credibility assessment: Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
Infusion phlebitis treatment	rate of resolution: recovery (treatment duration 3 days) VS 50% $MgSO_4$	Zheng,2014[(14)] (4)	AV	394	$<10^{-3}$	Y	Y (0.05)	Y (0.00)	N	N	High
Infusion phlebitis treatment	rate of resolution: recovery (treatment duration 15 days) VS 50% $MgSO_4$	Zheng,2014[(14)] (4)	AV	151	≤0.05	N	NA	Y (0.10)	NA	Y	High
Infusion phlebitis treatment	rate of resolution: marked improvement VS 50% $MgSO_4$	Zheng,2014[(14)] (9)	AV	814	$<10^{-3}$	Y	Y (0.01)	N (0.39)	N	Y	High
Infusion phlebitis treatment	rate of resolution: marked improvement (treatment duration 3 days) VS 50% $MgSO_4$	Zheng,2014[(14)] (7)	AV	679	≤0.05	Y	Y (0.03)	N (0.39)	N	Y	High
Infusion phlebitis treatment	rate of resolution: total improvement VS 50% $MgSO_4$	Zheng,2014[(14)] (10)	AV	880	$<10^{-6}$	N	Y (0.002)	N (0.98)	Y	Y	High

Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Criteria for credibility assessment: Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
Infusion phlebitis treatment	rate of resolution: total improvement (treatment duration 3 days) VS 50% $MgSO_4$	Zheng,2014[14] (8)	AV	679	$<10^{-6}$	N	Y (0.03)	N (1.00)	Y	Y	High
Infusion phlebitis treatment	rate of resolution: total improvement (treatment duration 15 days) VS 50% $MgSO_4$	Zheng,2014[14] (2)	AV	151	$<10^{-3}$	N	NA	N (0.14)	NA	Y	High
Infusion phlebitis treatment	rate of resolution: recovery VS non-AV medication	Zheng,2014[14] (3)	AV	283	≤0.05	Y	N (0.36)	N (0.33)	N	N	High
Infusion phlebitis treatment	rate of resolution: marked improvement VS non-AV medication	Zheng,2014[14] (3)	AV	163	≤0.05	N	N (0.65)	N (0.98)	N	N	High
Infusion phlebitis treatment	rate of resolution: total improvement VS non-AV medication	Zheng,2014[14] (4)	AV	323	≤0.05	N	N (0.92)	NP	N	Y	High

Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Criteria for credibility assessment: Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
OSF	burning sensation at 1 month	Al-Maweri, 2019[97](2)	AV	111	≤0.05	Y	NA	NP	NA	Y	Moderate
OSF	burning sensation at 2 months	Al-Maweri, 2019[97](2)	AV	111	$<10^{-3}$	N	NA	N (0.79)	NA	Y	Moderate
Glucose lowering	FBG in prediabetic & early nontreated DM	Zhang, 2016[12] (5)	AV	328	≤0.05	Y	N (0.28)	N (0.39)	N	Y	Low
Glucose lowering	HbA1C in prediabetic & early nontreated DM	Zhang, 2016[12] (2)	AV	76	$< 10^{-6}$	N	NA	NP	NA	Y	Low
Glucose lowering	FBG in type 2 DM	Suksomboon, 2016[93] (5)	AV	235	≤0.05	Y	N (0.71)	N (0.93)	N	Y	Low
Glucose lowering	HbA1C in type 2 DM	Suksomboon, 2016[93] (4)	AV	164	≤0.05	Y	N (0.77)	N (0.48)	N	Y	Low
Lipid lowering	TG	Zhang, 2016[12](4)	AV	206	$<10^{-3}$	Y	N (0.38)	NP	N	Y	Low
Lipid lowering	TC	Zhang, 2016[12](4)	AV	206	$<10^{-6}$	Y	N (0.31)	N (0.79)	N	Y	Low

Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Criteria for credibility assessment: Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
Lipid lowering	HDL	Zhang, 2016[12](3)	AV	136	≤0.05	Y	N (0.76)	N (1.00)	N	Y	Low
Lipid lowering	LDL	Zhang, 2016[12](3)	AV	136	$<10^{-6}$	Y	N (0.94)	N (0.71)	N	Y	Low
Irritate bowel syndrome	symptom scores improvement	Hong, 2018[11] (3)	AV	137	≤0.05	N	N (0.83)	NP	N	N	Critically low
Irritate bowel syndrome	short term symptom scores improvement (at 1 month)	Hong, 2018[11] (2)	AV	112	≤0.05	N	NA	NP	NA	N	Critically low
Irritate bowel syndrome	response rates	Hong, 2018[11] (2)	AV	112	≤0.05	N	NA	NP	NA	N	Critically low
CIP treatment	overall efficacy rate VS 50% $MgSO_4$	Gao, 2016[142] (6)	AV	547	$<10^{-6}$	N	N (0.56)	N (0.39)	Y	Y	Critically low
CIP treatment	overall cure rate VS 50% $MgSO_4$	Gao, 2016[142] (4)	AV	293	≤0.05	Y	N (0.28)	N (0.57)	N	Y	Critically low
Acute surgical wound	wound healing number	Wang, 2013[144] (2)	AV	97	$<10^{-3}$	N	NA	NA	NA	NR	Critically low

Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Criteria for credibility assessment: Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
Chronic wound	wound healing number	Wang, 2013[144] (5)	AV	233	≤0.05	Y	NA	NA	NA	NR	Critically low
Outcomes supported by non-significant evidence											
Infusion phlebitis prevention	total incidence (treatment duration 5 days) VS no treatment	Zheng, 2014[14](5)	AV	532	>0.05	Y	N (0.31)	N (0.61)	N	Y	High
Infusion phlebitis prevention	incidence of 3rd-degree phlebitis (treatment duration 5 days) VS no treatment	Zheng, 2014[14](4)	AV	482	>0.05	N	NA	NP	NA	N	High
Infusion phlebitis prevention	total incidence of phlebitis VS potato slice	Zheng, 2014[14](2)	AV	276	>0.05	N	NA	NP	NA	N	High
Infusion phlebitis prevention	total incidence of 2nd-degree phlebitis VS potato slice	Zheng, 2014[14](2)	C	276	>0.05	N	NA	NP	NA	N	High
Infusion phlebitis prevention	total incidence of phlebitis VS 50% $MgSO_4$	Zheng, 2014[14](2)	AV	248	>0.05	Y	NA	N (0.32)	NA	Y	High

Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Criteria for credibility assessment: Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
Radiation-induced reaction	mucositis prevention	Worthington, 2011[145](2)	AV	119	>0.05	Y	NA	NP	NA	Y	High
Burns	time to wound healing	Norman, 2017[143](3)	AV	210	>0.05	Y	NA	N (0.21)	NA	Y	High
Burns	infection	Norman, 2017[143](3)	AV	221	>0.05	N	NA	NP	NA	N	High
OSF	burning sensation at 3 months	Al-Maweri, 2019[97] (3)	AV	131	>0.05	Y	N (0.74)	N (0.77)	N	N	Moderate
OSF	mouth opening at 1 month	Al-Maweri, 2019[97] (5)	AV	319	>0.05	Y	Y (0.09)	N (1.00)	N	Y	Moderate
OSF	mouth opening at 2 months	Al-Maweri, 2019[97] (5)	AV	393	>0.05	Y	N (0.29)	N (0.99)	N	Y	Moderate
OSF	mouth opening at 3 months	Al-Maweri, 2019[97] (6)	AV	413	>0.05	Y	N (0.28)	N (1.00)	N	Y	Moderate
OSF	tongue protrusion at 1 month	Al-Maweri, 2019[97] (4)	AV	351	>0.05	Y	N (0.62)	N (0.14)	N	N	Moderate

Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Criteria for credibility assessment: Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
OSF	tongue protrusion at 2 months	Al-Maweri, 2019[97] (4)	AV	351	>0.05	Y	N (0.54)	NP	N	Y	Moderate
OSF	tongue protrusion at 3 months	Al-Maweri, 2019[97] (5)	C	371	>0.05	Y	N (0.60)	N (0.33)	N	N	Moderate
OSF	cheek flexibility at 1 month	Al-Maweri, 2019[97] (2)	AV	111	>0.05	Y	NA	NP	NA	Y	Moderate
OSF	cheek flexibility at 2 months	Al-Maweri, 2019[97] (2)	AV	111	>0.05	Y	NA	NP	NA	Y	Moderate
OSF	cheek flexibility at 3 months	Al-Maweri, 2019[97] (3)	AV	131	>0.05	Y	N (0.64)	NP	N	Y	Moderate
Glucose lowering	insulin level in prediabetic & early nontreated DM	Zhang, 2016[12](2)	AV	151	>0.05	Y	NA	NP	NA	Y	Low
OLP	pain and burning sensation	Ali, 2017[98](3)	AV	121	>0.05	N	Y (0.06)	N (0.97)	N	Y	Critically low
OLP	clinical improvement	Ali, 2017[98](3)	AV	121	>0.05	N	N (0.30)	N (1.00)	N	Y	Critically low

Indication	Outcome examined	Author, Year[Ref] (No.of study included)	Results favor	Criteria for credibility assessment: Total n	*P* value random effects	I^2 >50%	SSE (p-value)[a]	ESB (p-value)[b]	95% PI exclude null	LS	AMSTAR quality
Irritate bowel syndrome	long-term symptoms scores improvement (at 3 months)	Hong, 2018[(11)](2)	AV	67	>0.05	N	NA	NP	NA	N	Critically low
CIP prevention	incidence of 1st-degree CIP	Gao, 2016[(142)](10)	AV	3925	>0.05	Y	N (0.48)	NP	N	N	Critically low
Acute surgical wound	mean time to wound healing	Wang, 2013[(144)](2)	AV	101	>0.05	Y	NA	NA	NA	N R	Critically low

[a] p-value from Egger test, significant threshold p ≤ 0.1.

[b] significant threshold p≤0.1

Abbreviations: Ref – reference number, AV – *Aloe vera,* n – number of participants, , I^2 – heterogeneity, SSE – small study effects, ESB – excess significance, PI – prediction interval, LS – largest study showed a statistically significance, AMSTAR – A Measurement Tool to Assess Systematic Reviews, VS – versus, C – control, Y-yes, N – no, , NA– not applicable, NP – not pertinent because of fewer-than-expected number of observed studies, CIP – chemotherapy induced phlebitis, $MgSO_4$ – magnesium sulfate, OSF – oral submucous fibrosis, OLP – oral lichen planus, IBS – irritate bowel syndrome, FBG – fasting blood glucose, HbA1C – Hemoglobin A1c, TC – total cholesterol, TG – triglyceride, LDL – low-density lipoprotein cholesterol, HDL – high- density lipoprotein cholesterol, DM – diabetes mellitus

Table 11 Methodological quality assessment using AMSTAR 2 instrument

Author,Year	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Rating
Ali,2017 (98)	Y	N	Y	PY	Y	Y	Y	PY	Y	Y	Y	Y	Y	Y	N	Y	Critically low
Al-Maweri(a), 2019(97)	Y	PY	Y	PY	Y	N	Y	PY	Y	Y	Y	N	Y	Y	Y	Y	Moderate
Gao,2016 (142)	Y	N	Y	PY	N	Y	N	PY	PY	N	Y	Y	Y	Y	Y	Y	Critically low
Hong,2018 (11)	Y	N	Y	PY	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Critically low
Norman,2017 (143)	Y	PY	Y	PY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	High
Suksomboon,2016 (93)	Y	N	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Low
Wang, 2013 (144)	Y	N	Y	PY	Y	Y	N	PY	Y	N	Y	Y	Y	N	Y	N	Critically low
Worthington,2011 (145)	Y	Y	Y	PY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	High
Zhang,2016 (12)	Y	N	Y	PY	Y	Y	N	Y	PY	N	Y	N	Y	Y	Y	Y	Low
Zheng,2014 (14)	Y	PY	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	High

Abbreviations: Q – question, Y – yes, N – no, PY – partial yes

Q1. Did the research questions and inclusion criteria for the review include the components of PICO?

Q2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

Q3. Did the review authors explain their selection of the study designs for inclusion in the review?

Q4. Did the review authors use a comprehensive literature search strategy?

Q5. Did the review authors perform study selection in duplicate?

Q6. Did the review authors perform data extraction in duplicate?

Q7. Did the review authors provide a list of excluded studies and justify the exclusions?

Q8. Did the review authors describe the included studies in adequate detail?

Q9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

Q10. Did the review authors report on the sources of funding for the studies included in the review?

Q11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?

Q12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?

Q13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?

Q14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

Q15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?

Q16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

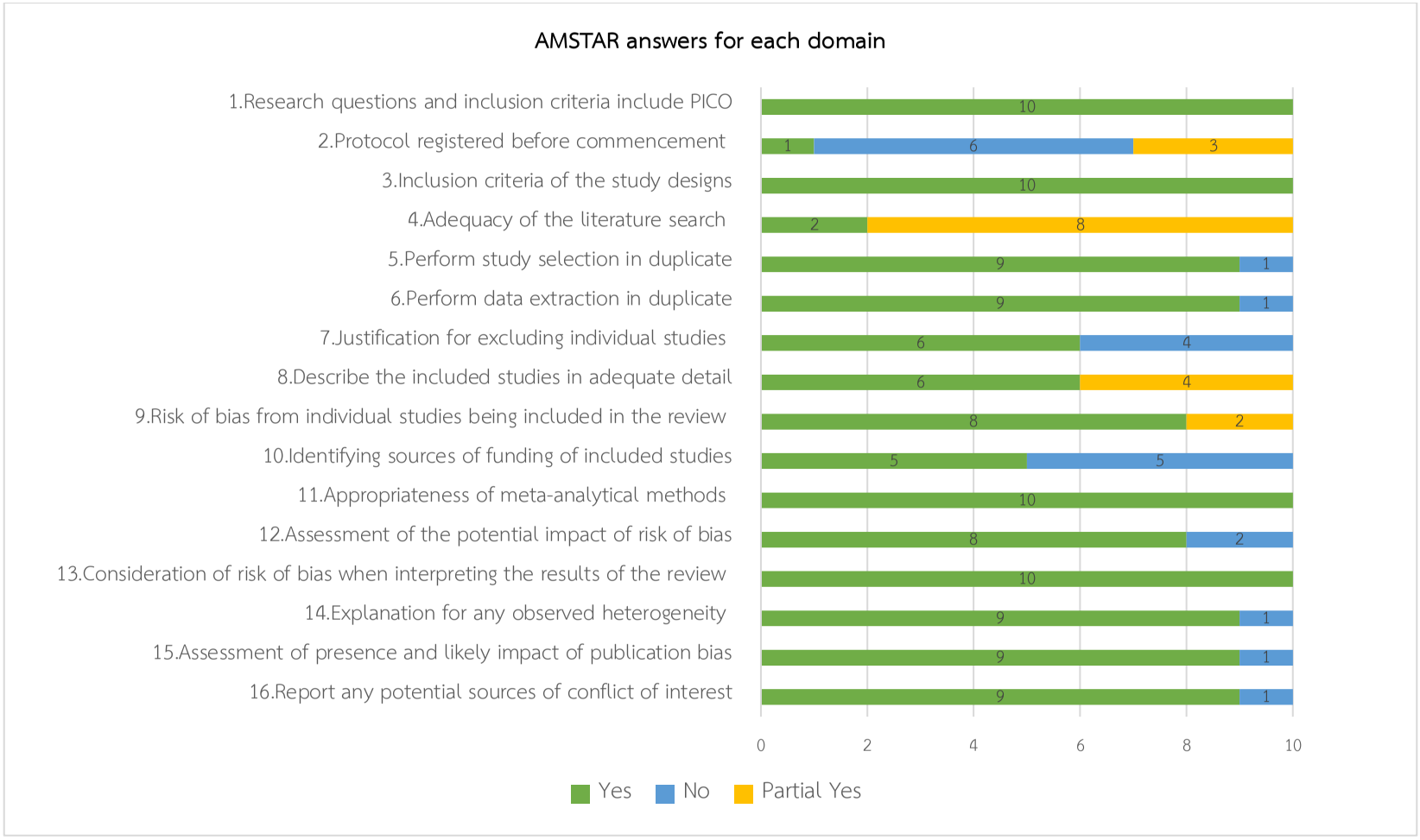

Figure 4 Answers of each domain in AMSTAR 2

7. Sensitivity analysis

When sensitivity analyses of RCTs were performed, the evidence was being upgraded from highly suggestive to convincing evidence in 1 outcome examined (prevention of second-degree phlebitis induced by intravenous infusion when compared with no treatment). Meta-analysis that examined effect of *Aloe vera* in the prevention of third-degree infusion phlebitis compared with no treatment retained the same rank as suggestive evidence. Other outcomes that examined the effect of *Aloe vera* in the prevention of CIP were downgraded to weak evidence, as shown in Table 12.

Table 12 Sensitivity analysis of only RCTs included in the evidence that graded as highly suggestive or suggestive evidence (n=5)

					Criteria for credibility assessment							
Indication	Outcome examined	No.of included study	Author, Year	Effect size (95% CI)[a]	Total n	P-value for random effects	I^2 >50%	SSE (P-value)[b]	ESB	95% PI	LS	Change of level of evidence
infusion phlebitis prevention	incidence of 2nd-degree phlebitis (VS no treatment)	10	Zheng, 2014 (14)	RR 0.20 (0.12-0.33)	1119	$4.62x10^{-10}$	25.9	N (0.65)	NP	0.06 to 0.59	Y	Highly suggestive to convincing
infusion phlebitis prevention	incidence of 3rd-degree phlebitis (VS no treatment)	8	Zheng, 2014 (14)	RR 0.18 (0.08-0.43)	1119	1.22×10^{-4}	0.0	N (0.48)	NP	0.06 to 0.54	N	Suggestive retained
CIP prevention	incidence of 2nd-degree CIP	3	Gao,2016 (142)	OR 0.11 (0.05-0.25)	416	9.18×10^{-8}	0.0	N (0.56)	NP	0.0 to 19.96	Y	Highly suggestive to weak
CIP prevention	overall incidence	4	Gao,2016 (142)	OR 0.14 (0.05-0.37)	408	9.9×10^{-5}	72.1	N (0.29)	NP	0.00 to 10.33	Y	Highly suggestive to weak
CIP prevention	incidence of 3rd-degree CIP	3	Gao,2016 (142)	OR 0.14 (0.04-0.54)	416	0.004	0.0	N (0.93)	NP	0.0 to 891.52	N	Suggestive to weak

[a] Effect size based on random-effects model

[b] significant threshold $p \leq 0.1$

Abbreviations: CI – confidence interval, n – number of participants, I^2 – heterogeneity, SSE – small study effects, ESB – excess significance bias, PI – prediction interval, LS – largest study showed a statistically significance, VS – versus, RR – relative risk, OR – odds ratio, N – no, Y – yes, NP – not pertinent because of fewer-than-expected number of observed studies, CIP – chemotherapy induced phlebitis

CHAPTER V

DISCUSSION AND CONCLUSIONS

1. Discussion

To our knowledge, this is the first umbrella review of systematic reviews and meta-analyses of clinical trials that evaluated the effect of *Aloe vera* on health outcomes. Overall, 10 systematic reviews and 10 meta-analyses with 71 unique outcomes of *Aloe vera* have been considered. Using criteria for credibility assessment, none of evidence reached the convincing level. Only 3 highly suggestive evidence supported benefits of *Aloe vera* in the prevention of second-degree infusion phlebitis relative to no treatment and prevention of chemotherapy-induced phlebitis (CIP) regarding the reduction of overall incidence and incidence of the second degree of severity. Two suggestive evidence supported benefits of *Aloe vera* in prevention of third-degree infusion phlebitis when compared with no treatment and prevention of the third-degree CIP when compared with conventional treatment. A sensitivity analysis limited to RCTs showed that effect of *Aloe vera* in the prevention of second-degree infusion phlebitis was being upgraded to having convincing evidence and the prevention of third-degree infusion phlebitis remained in suggestive level. However, the others were downgraded into weak level. Overall, the results showed that most of the effect of *Aloe vera* on health outcomes was not supported by high-level-of credibility evidence.

Phlebitis is an inflammation of the vein caused by chemical, mechanical, or infectious irritation (146-148). Several pharmacological interventions (e.g., non-steroidal anti-inflammatory drugs (NSAIDs) (149), heparin (150, 151), steroid ointment (152), and traditional medicines such as *Sesame indicum* (153), nigella *sativa* (154), and potato (155)) have been suggested to help reduce incidence of infusion phlebitis and CIP. However, the number of evidences is limited and it is yet unknown what are

the most efficient methods. To date, there is no strong recommendation for using any medication for phlebitis prevention. This study found that *Aloe vera* could prevent phlebitis induced by chemotherapy and intravenous infusion. The possible mechanism has been suggested that *Aloe vera* had healing properties, anti-inflammatory activity, effects on the immune system, skin protection, and antiseptic effects (26, 31). Fresh *Aloe vera* has been found to promote the attachment and increase the healing of wounded monolayer of cells whereby *Aloe vera* gel enhanced the content of collagen and degree of collagen cross linking (4). Thus, *Aloe vera* might be beneficial for the prevention and treatment of phlebitis.

Considering the results of this umbrella review, *Aloe vera* showed promising results in the prevention of second and third-degree phlebitis induced by intravenous access with highly suggestive and suggestive evidence. Despite no small-study effects in these meta-analyses, large heterogeneity was reported. Additionally, benefits of *Aloe vera* in the prevention of CIP regarding the reduction of overall incidence, incidence of the second and third-degree CIP were supported by highly suggestive and suggestive evidence. Among these studies, the highly suggestive and suggestive credibility level was expressed by large sample size, a p-value less than 10^{-3}, and no large heterogeneity was found in the meta-analysis examined the prevention of second-degree CIP. However, summary effect sizes were not relatively large and small-study effects were evident. The methodological quality based on AMSTAR 2 assessment was rated as high for the meta-analysis examined effect of *Aloe vera* in the infusion phlebitis prevention, apart from critically low in meta-analysis examined the CIP prevention. Therefore, the results need to be interpreted with caution.

Regarding positive results of *Aloe vera* in the phlebitis prevention and treatment, *Aloe vera* gel or leaves were used for external application in these meta-analyses without chemically treated, which were different from other meta-analyses

included in this umbrella review. Using fresh *Aloe vera* instead of the *Aloe vera*-derived preparations such as gel, cream, or ointment, might be inconvenient and percentage of active ingredients might vary. However, considering the high incidence of phlebitis induced by intravenous injection and chemotherapy drugs, risk of developing serious complications, and the potential additional treatment costs, the results of current study should be implemented (156-158). *Aloe vera* should be suggested as an effective complementary alternative medicine for the prevention of phlebitis, particularly in high degree of severity.

Large proportion of outcomes (59%) included in this umbrella review were supported by the evidence with a weak level of credibility. These outcomes included the effect of *Aloe vera* in symptoms improvement for irritable bowel syndrome, which is the widely used indication of *Aloe vera* (7). In addition, *Aloe vera* also showed positive effects in reduction of time to healing in acute-surgical and chronic wounds, reduction of a burning sensation among patients with oral submucous fibrosis for 2 months, reduction of FBG, HBA_1C in prediabetic and early non-treated diabetic patients, and reduction of TG, TC, and LDL level and increment of the HDL levels in type 2 diabetic patients. All of these evidences were statistically significant suggests positive effect of *Aloe vera* and some of them were reported with large effect sizes, however, were graded as a weak level of credibility of evidence due to small sample size and some of these outcomes also had high heterogeneity. For these reasons, implementation of *Aloe vera* in these health outcomes in clinical practice should be done with caution. Moreover, we found that the effect of *Aloe vera* on burn wound treatment, the well-known indication (9, 159), was non-significant. Therefore, further studies are needed to confirm the effects of *Aloe vera*.

High heterogeneity was detected in most of the included meta-analyses. This is probably caused by different types of *Aloe vera* used as described earlier. The majority of the included studies did not consider the amounts of active ingredients

which may affect the therapeutic effects of *Aloe vera*. The amounts of active ingredients can vary among *Aloe vera* preparations, depending on harvesting and storage conditions, parts of plants used, the time of used after harvesting, and extraction methods (31, 53). Furthermore, variability in study design may cause heterogeneity (160). Sensitivity analysis in this study suggested that limiting only RCTs could reduce a degree of heterogeneity and also upgrade the evidence in prevention of second-degree infusion phlebitis from highly suggestive to having convincing evidence. Moreover, variation in co-intervention and compliance may have an important role. Most of the included meta-analyses did not report on the patient's compliance. Some of the systematic reviews and meta-analyses used the combination of *Aloe vera* with other medications or herbal medicines, making it difficult to determine the true effect of *Aloe vera*. On the other hand, results of this umbrella review are more generalizable because such combination was generally found in real-world practice.

The strengths of this study include using data from systematic reviews and meta-analyses of clinical trials, the appropriate study designs to investigate the effect of the given intervention. Furthermore, this umbrella review incorporating articles without language restriction, which would cover all related studies available in this field. However, findings from this study had some limitations. First, various meta-analyses pooled a small number of studies, leading to the risk for small-study effects. Second, the quality of the individual component primary studies was not appraised in this study because this was beyond the scope of umbrella review. Third, this study assessed only data from previously published systematic reviews and meta-analyses. Thus, other information that have not yet been published and the primary studies that have not yet been assessed through meta-analytic approaches might have been missed. Additionally, despite the use of *Aloe vera* in different doses, preparation, and

dosage regimen, the included studies did not consider these factors which may affect the outcomes of *Aloe vera*. Thus, further investigation of these factors in future studies are needed. Finally, long-term benefit remains to be determined due to the findings of this reviews showed that the longest duration of *Aloe vera* used were 3 months (12). Regarding the evidence of carcinogenic activity in animal model, the long-term safety also needs to be concerned (29).

2. Conclusions

In summary, this umbrella review of the effects of *Aloe vera* on health outcomes found that the current suggestive evidence suggests the benefits of *Aloe vera* in the prevention of phlebitis induced by chemotherapy and intravenous infusion, particularly in severe stage. Nevertheless, most of the current evidence had limitations including poor methodological quality and small number of participants included. The benefit of *Aloe vera* should therefore be reviewed with caution and data from more well-designed, larger number of participants, and robust studies using standardized preparations are needed to confirm the benefit of *Aloe vera* on health outcomes.

REFERENCES

1. Barnes PM, Bloom B, Nahin RL. Complementary and alternative medicine use among adults and children: United States, 2007. Natl Health Stat Report. 2008(12):1-23.

2. Kemppainen LM, Kemppainen TT, Reippainen JA, Salmenniemi ST, Vuolanto PH. Use of complementary and alternative medicine in Europe: Health-related and sociodemographic determinants. Scand J Public Health. 2018;46(4):448-55.

3. Peltzer K, Pengpid S. The use of herbal medicines among chronic disease patients in Thailand: a cross-sectional survey. J Multidiscip Healthc. 2019;12:573-82.

4. Surjushe A, Vasani R, Saple DG. Aloe vera: a short review. Indian J Dermatol. 2008;53(4):163-6.

5. Samarh SN, Khalaf NA, Hajhamad MM. Evidence based medical use of aloe vera extracts, short review of literature. Int J Res Med Sci. 2017;5(10):4198-202.

6. Nutrition business journal. An analysis of markets, trends, competition and strategy in the U.S. dietary supplement industry 2012 [cited 2019 31 October]. Available from: http://newhope360.com/2012-supplement-business-report.

7. European Medicines Agency. European Union herbal monograph on Aloe barbadensis Mill. and on Aloe (various species, mainly Aloe ferox Mill. and its hybrids), folii succus siccatus London, United Kingdom: European Medicines Agency; 2017 [cited 2019 October 31]. Available from: https://www.ema.europa.eu/en/documents/herbal-monograph/final-european-union-herbal-monograph-aloe-barbadensis-mill-aloe-various-species-mainly-aloe-ferox_en.pdf.

8. National center of complementary and integrative health. Aloe vera [updated September 2016 cited 2019 October 31]. Available from: https://nccih.nih.gov/health/aloevera

9. World Health Organization. WHO monographs on selected medicinal plants.-Vol. 1. Malta1999 [cited 2019 November 3]. Available from: https://apps.who.int/medicinedocs/pdf/s2200e/s2200e.pdf.

10. Maenthaisong R, Chaiyakunapruk N, Niruntraporn S, Kongkaew C. The efficacy of aloe vera used for burn wound healing: a systematic review. Burns. 2007;33(6):713-8.

11. Hong SW, Chun J, Park S, Lee HJ, Im JP, Kim JS. Aloe vera is effective and safe in short-term treatment of irritable bowel syndrome: a systematic review and meta-analysis. J Neurogastroenterol Motil. 2018;24(4):528-35.

12. Zhang Y, Liu W, Liu D, Zhao T, Tian H. Efficacy of Aloe vera supplementation on prediabetes and early non-treated diabetic patients: a systematic review and meta-analysis of randomized controlled trials. Nutrients. 2016;8(7):388.

13. Suksomboon N, Poolsup N, Punthanitisarn S. Effect of Aloe vera on glycaemic control in prediabetes and type 2 diabetes: A systematic review and meta-analysis. Journal of Clinical Pharmacy and Therapeutics. 2016;41(2):180-8.

14. Zheng GH, Yang L, Chen HY, Chu JF, Mei L. Aloe vera for prevention and treatment of infusion phlebitis. Cochrane Database Syst Rev. 2014(6):CD009162.

15. Fusar-Poli P, Radua J. Ten simple rules for conducting umbrella reviews. Evid Based Mental Health. 2018;21(3):95-100.

16. Ioannidis JP. Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. CMAJ. 2009;181(8):488-93.

17. Biondi-Zoccai G. Umbrella reviews. Switzerland: Springer International; 2016.

18. Belbasis L, Bellou V, Evangelou E, Ioannidis JP, Tzoulaki I. Environmental risk factors and multiple sclerosis: an umbrella review of systematic reviews and meta-analyses. Lancet Neurol. 2015;14(3):263-73.

19. Kyrgiou M, Kalliala I, Markozannes G, Gunter MJ, Paraskevaidis E, Gabra H, et al. Adiposity and cancer at major anatomical sites: umbrella review of the literature. Bmj. 2017;356:j477.

20. Kalliala I, Markozannes G, Gunter MJ, Paraskevaidis E, Gabra H, Mitra A, et al. Obesity and gynaecological and obstetric conditions: umbrella review of the literature. Bmj. 2017;359:j4511.

21. Dragioti E, Evangelou E, Larsson B, Gerdle B. Effectiveness of multidisciplinary programmes for clinical pain conditions: an umbrella review. J Rehabil Med. 2018;50(9):779-91.

22. Dragioti E, Solmi M, Favaro A, Fusar-Poli P, Dazzan P, Thompson T, et al. Association of Antidepressant Use With Adverse Health Outcomes: A Systematic

Umbrella Review. JAMA Psychiatry. 2019;76(12):1241-55.

23. Chakranon P, Lai YK, Tang YW, Choudhary P, Khunti K, Lee SWH. Distal technology interventions in people with diabetes: an umbrella review of multiple health outcomes. Diabet Med. 2019;00:1-11.

24. Wan Q, Li N, Du L, Zhao R, Yi M, Xu Q, et al. Allium vegetable consumption and health: an umbrella review of meta-analyses of multiple health outcomes. Food Sci Nutr. 2019;7(8):2451-70.

25. Dinu M, Pagliai G, Casini A, Sofi F. Mediterranean diet and multiple health outcomes: an umbrella review of meta-analyses of observational studies and randomised trials. Eur J Clin Nutr. 2018;72(1):30-43.

26. Radha MH, Laxmipriya NP. Evaluation of biological properties and clinical effectiveness of Aloe vera: A systematic review. J Tradit Complement Med. 2015;5(1):21-6.

27. Sharma P, Kharkwal A, Kharkwal H, Abdin M, Varma A. A review on pharmacological properties of Aloe vera. Int J Pharm Sci Rev Res. 2014;29:31-7.

28. Salehi B, Albayrak S, Antolak H, Krȩgiel D, Pawlikowska E, Sharifi-Rad M, et al. Aloe Genus Plants: From Farm to Food Applications and Phytopharmacotherapy. Int J Mol Sci. 2018;19(9):E2843.

29. Guo X, Mei N. Aloe vera: A review of toxicity and adverse clinical effects. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev. 2016;34(2):77-96.

30. Cock I. The genus aloe: phytochemistry and therapeutic uses including treatments for gastrointestinal conditions and chronic inflammation. Novel natural products: therapeutic effects in pain, arthritis and gastro-intestinal diseases: Springer; 2015. p. 179-235.

31. Boudreau MD, Beland FA. An evaluation of the biological and toxicological properties of Aloe barbadensis (miller), Aloe vera. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev. 2006;24(1):103-54.

32. Chantarawaratit P, Sangvanich P, Banlunara W, Soontornvipart K, Thunyakitpisal P. Acemannan sponges stimulate alveolar bone, cementum and periodontal ligament regeneration in a canine class II furcation defect model. J Periodontal Res.

2014;49(2):164-78.

33. Rajput SS, Soni KK, Saxena R. Pharmacology and phytochemistry of saponin isolated from Aloe vera for wound healing activity. Asian J Chem. 2009;21(2):1029-32.

34. Choi SW, Son BW, Son YS, Park YI, Lee SK, Chung MH. The wound-healing effect of a glycoprotein fraction isolated from aloe vera. Br J Dermatol. 2001;145(4):535-45.

35. Chithra P, Sajithlal GB, Chandrakasan G. Influence of aloe vera on the healing of dermal wounds in diabetic rats. J Ethnopharmacol. 1998;59(3):195-201.

36. Rodriguez-Bigas M, Cruz NI, Suarez A. Comparative evaluation of Aloe vera in the management of burn wounds in Guinea pigs. Plast Reconstr Surg. 1988;81(3):386-9.

37. Khorasani G, Hosseinimehr SJ, Azadbakht M, Zamani A, Mahdavi MR. Aloe versus silver sulfadiazine creams for second-degree burns: a randomized controlled study. Surg Today. 2009;39(7):587-91.

38. Visuthikosol V, Chowchuen B, Sukwanarat Y, Sriurairatana S, Boonpucknavig V. Effect of aloe vera gel to healing of burn wound a clinical and histologic study. J Med Assoc Thai. 1995;78(8):403-9.

39. Shahzad MN, Ahmed N. Effectiveness of Aloe Vera gel compared with 1% silver sulphadiazine cream as burn wound dressing in second degree burns. J Pak Med Assoc. 2013;63(2):225-30.

40. Burusapat C, Supawan M, Pruksapong C, Pitiseree A, Suwantemee C. Topical aloe vera gel for accelerated wound healing of split-thickness skin graft donor sites: a double-blind, randomized, controlled trial and systematic review. Plast Reconstr Surg. 2018;142(1):217-26.

41. Puvabanditsin P, Vongtongsri R. Efficacy of aloe vera cream in prevention and treatment of sunburn and suntan. J Med Assoc Thai. 2005;88 Suppl 4:S173-6.

42. Khorasani G, Ahmadi A, Jalal Hosseinimehr S, Ahmadi A, Taheri A, Fathi H. The effects of Aloe vera cream on split-thickness skin graft donor site management: a randomized, blinded, placebo-controlled study. Wounds. 2011;23(2):44.

43. Tanaka M, Misawa E, Ito Y, Habara N, Nomaguchi K, Yamada M, et al. Identification of five phytosterols from Aloe vera gel as anti-diabetic compounds. Biol Pharm Bull. 2006;29(7):1418-22.

44. Anand S, Muthusamy VS, Sujatha S, Sangeetha KN, Bharathi Raja R, Sudhagar S, et al. Aloe emodin glycosides stimulates glucose transport and glycogen storage through PI3K dependent mechanism in L6 myotubes and inhibits adipocyte differentiation in 3T3L1 adipocytes. FEBS Letters. 2010;584(14):3170-8.

45. Shin E, Shin S, Kong H, Lee S, Do SG, Jo TH, et al. Dietary Aloe reduces adipogenesis via the activation of AMPK and suppresses obesity-related Inflammation in obese mice. Immune Netw. 2011;11(2):107-13.

46. Beppu H, Nagamura Y, Fujita K. Hypoglycaemic and antidiabetic effects in mice of Aloe arborescens Miller var. natalensis Berger. Phytother Res. 1993;7(7):S37-42.

47. Rajasekaran S, Ravi K, Sivagnanam K, Subramanian S. Beneficial effects of aloe vera leaf gel extract on lipid profile status in rats with streptozotocin diabetes. Clin Exp Pharmacol Physiol. 2006;33(3):232-7.

48. Bunyapraphatsara N, Yongchaiyudha S, Rungpitarangsi V, Chokechaijaroenporn O. Antidiabetic activity of Aloe vera L. juice II. Clinical trial in diabetes mellitus patients in combination with glibenclamide. Phytomedicine. 1996;3(3):245-8.

49. Yongchaiyudha S, Rungpitarangsi V, Bunyapraphatsara N, Chokechaijaroenporn O. Antidiabetic activity of Aloe vera L. juice. I. Clinical trial in new cases of diabetes mellitus. Phytomedicine. 1996;3(3):241-3.

50. Alinejad-Mofrad S, Foadoddini M, Saadatjoo SA, Shayesteh M. Improvement of glucose and lipid profile status with Aloe vera in pre-diabetic subjects: a randomized controlled-trial. J Diabetes Metab Disord. 2015;14:22.

51. Ulbricht C, Armstrong J, Basch E, Basch S, Bent S, Dacey C, et al. An evidence-based systematic review of aloe vera by the natural standard research collaboration. J Herb Pharmacother. 2007;7(3-4):279-323.

52. Foster M HD, Samman S. Evaluation of the nutritional and metabolic effects of Aloe vera. In: Benzie IFF W-GS, editor. Herbal Medicine: Biomolecular and Clinical Aspects. 2nd edition ed. Boca Raton, FL: CRC Press/Taylor & Francis; 2011.

53. Hamman JH. Composition and applications of Aloe vera leaf gel. Molecules. 2008;13(8):1599-616.

54. Borra SK, Lagisetty RK, Mallela GR. Anti-ulcer effect of Aloe vera in non-steroidal anti-inflammatory drug induced peptic ulcers in rats. Afr J Pharm Pharmacol

2011;5(16):1867-71.

55. Eamlamnam K, Patumraj S, Visedopas N, Thong-Ngam D. Effects of Aloe vera and sucralfate on gastric microcirculatory changes, cytokine levels and gastric ulcer healing in rats. World J Gastroenterol. 2006;12(13):2034-9.

56. Bautista-Pérez R, Segura-Cobos D, Vázquez-Cruz B. In vitro antibradykinin activity of Aloe barbadensis gel. J Ethnopharmacol. 2004;93(1):89-92.

57. Davis RH, Donato J, Hartman GM, Haas RC. Anti-inflammatory and wound healing activity of a growth substance in Aloe vera. J Am Podiatr Med Assoc. 1994;84(2):77-81.

58. Alemdar S, Agaoglu S. Investigation of in vitro antimicrobial activity of Aloe vera juice. J Anim Vet Adv. 2009;8(1):99-102.

59. Philip J, John S, Iyer P. Antimicrobial activity of aloe vera barbedensis, daucus carota, emblica officinalis, honey and punica granatum and formulation of a health drink and salad. Mal J Microbiol. 2012;8(3):141-7.

60. Olaleye M, Bello-Michael C. Comparative antimicrobial activities of Aloe vera gel and leaf. Afr J Biotechnol. 2005;4(12):1413-4.

61. Saoo K, Miki H, Ohmori M, Winters W. Antiviral activity of aloe extracts against cytomegalovirus. Phytother Res. 1996;10(4):348-50.

62. Sydiskis R, Owen D, Lohr J, Rosler K, Blomster R. Inactivation of enveloped viruses by anthraquinones extracted from plants. Antimicrob Agents Chemother. 1991;35(12):2463-6.

63. Li S-W, Yang T-C, Lai C-C, Huang S-H, Liao J-M, Wan L, et al. Antiviral activity of aloe-emodin against influenza A virus via galectin-3 up-regulation. Eur J Pharmacol. 2014;738:125-32.

64. Rahmani AH, Aldebasi YH, Srikar S, Khan AA, Aly SM. Aloe vera: potential candidate in health management via modulation of biological activities. Pharmacogn Rev. 2015;9(18):120-6.

65. El-Soudany K, Yagi A, Kabbash A. A self-controlled single blinded clinical trial to evaluate oral lichen planus after topical treatment with Aloe vera. J Gastroenterol Hepatol Res. 2013;2(4):503-7.

66. Babaee N, Zabihi E, Mohseni S, Moghadamnia AA. Evaluation of the therapeutic effects of Aloe vera gel on minor recurrent aphthous stomatitis. Dent Res J (Isfahan).

2012;9(4):381-5.

67. Sudarshan R, Annigeri RG, Sree Vijayabala G. Aloe vera in the treatment for oral submucous fibrosis–a preliminary study. J Oral Pathol Med. 2012;41(10):755-61.

68. West DP, Zhu YF. Evaluation of aloe vera gel gloves in the treatment of dry skin associated with occupational exposure. Am J Infect Control. 2003;31(1):40-2.

69. Di Franco R, Sammarco E, Calvanese MG, De Natale F, Falivene S, DiLecce A, et al. Preventing the acute skin side effects in patients treated with radiotherapy for breast cancer: the use of corneometry in order to evaluate the protective effect of moisturizing creams. Radiat Oncol. 2013;8(1):57.

70. Lee CK, Han SS, Shin YK, Chung MH, Park YI, Lee SK, et al. Prevention of ultraviolet radiation-induced suppression of contact hypersensitivity by Aloe vera gel components. Int J Immunopharmacol. 1999;21(5):303-10.

71. Dat AD, Poon F, Pham KB, Doust J. Aloe vera for treating acute and chronic wounds. Cochrane Database Syst Rev. 2012;15(2):CD008762.

72. Djuv A, Nilsen OG. Aloe vera juice: IC50 and dual mechanistic inhibition of CYP3A4 and CYP2D6. Phytother Res. 2012;26(3):445-51.

73. Lee A, Chui PT, Aun CS, Gin T, Lau AS. Possible interaction between sevoflurane and Aloe vera. Ann Pharmacother. 2004;38(10):1651-4.

74. Grindlay D, Reynolds T. The aloe vera phenomenon: a review of the properties and modern uses of the leaf parenchyma gel. J Ethnopharmacol. 1986;16(2-3):117-51.

75. Yaron A. Characterization of Aloe vera gel before and after autodegradation, and stabilization of the natural fresh gel. Phytother Res. 1993;7(7):S11-3.

76. Pellizzoni M, Molinari GP, Lucini L. Stability of the main Aloe fractions and Aloe-based commercial products under different storage conditions. Agrochimica. 2011;55(5):288-96.

77. Morrow DM, Rapaport MJ, Strick RA. Hypersensitivity to aloe. Arch Dermatol. 1980;116(9):1064-5.

78. Rabe C, Musch A, Schirmacher P, Kruis W, Hoffmann R. Acute hepatitis induced by an Aloe vera preparation: a case report. World J Gastroenterol. 2005;11(2):303-4.

79. Kanat O, Ozet A, Ataergin S. Aloe vera-induced acute toxic hepatitis in a healthy young man. Eur J Intern Med. 2006;17(8):589.

80. World Health Organization. International classification of impairments, disabilities, and handicaps : a manual of classification relating to the consequences of disease, published in accordance with resolution WHA29.35 of the Twenty-ninth World Health Assembly, May 1976 Geneva: World Health Organization; 1980 [cited 2019 1 December]. Available from: https://apps.who.int/iris/handle/10665/41003.

81. Ronen GM, Rosenbaum PL. Health outcomes measurement: concepts, guidelines and opportunities. In: Stone TE, editor. Handbook of clinical neurology. 111. Amsterdam, The Netherlands: Elsevier; 2013. p. 35-41.

82. Velentgas P, Dreyer NA, Wu AW. Outcome definition and measurement. In: Velentgas P DN, Nourjah P, et al., editor. Developing a protocol for observational comparative effectiveness research: a user's guide. Rockville, MD: Agency for Healthcare Research and Quality (US); 2013. p. 71-92.

83. Kelley TA. International Consortium for Health Outcomes Measurement (ICHOM). Trials. 2015;16(Suppl 3):O4.

84. Canadian institute for health information. Outcomes 2019 [cited 2019 26 October]. Available from: https://www.cihi.ca/en/outcomes.

85. Sansoni JE, editor Health outcomes: an overview from an Australian perspective. Australian health outcomes collaboration; 2016 23 August 2016; Australian health services research institute, University of Wollongong, Australia.

86. Oleske DM, Islam SS. Role of Epidemiology in the biopharmaceutical industry. In: Doan T, Renz C, Bhattacharya M, Lievano F, Scarazzini L, editors. Pharmacovigilance: A Practical Approach: Elsevier; 2019. p. 69-87.

87. Guyatt G, Rennie D, Meade M, Cook D. Users' guides to the medical literature: a manual for evidence-based clinical practice. 3 ed. Chicago, IL: AMA press Chicago; 2002.

88. Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions version 6.0: Cochrane; 2019 [updated July 2019; cited 2019 20 October]. Available from: www.training.cochrane.org/handbook.

89. Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. Health Info Libr J. 2009;26(2):91-108.

90. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that

evaluate health care interventions: explanation and elaboration. PLoS Med. 2009;6(7):e1000100.

91. Zhang Y, Liu W, Liu D, Zhao T, Tian H. Efficacy of aloe vera supplementation on prediabetes and early non-treated diabetic patients: A systematic review and meta-analysis of randomized controlled trials. Nutrients. 2016;8(7):E388.

92. Dick WR, Fletcher EA, Shah SA. Reduction of fasting blood glucose and hemoglobin A1C using oral aloe vera: a meta-analysis. J Altern Complement Med. 2016;22(6):450-7.

93. Suksomboon N, Poolsup N, Punthanitisarn S. Effect of aloe vera on glycaemic control in prediabetes and type 2 diabetes: a systematic review and meta-analysis. J Clin Pharm Ther. 2016;41(2):180-8.

94. Langhorst J, Wulfert H, Lauche R, Klose P, Cramer H, Dobos GJ, et al. Systematic review of complementary and alternative medicine treatments in inflammatory bowel diseases. J Crohns Colitis. 2015;9(1):86-106.

95. Nair GR, Naidu GS, Jain S, Nagi R, Makkad RS, Jha A. Clinical effectiveness of aloe vera in the management of oral mucosal diseases-a systematic review. J Clin Diagn Res. 2016;10(8):ZE01-7.

96. Mangaiyarkarasi SP, Manigandan T, Elumalai M, Cholan PK, Kaur RP. Benefits of Aloe vera in dentistry. J Pharm Bioallied Sci. 2015;7(Suppl 1):S255-9.

97. Al-Maweri SA, Ashraf S, Lingam AS, Alqutaibi A, Abdulrab S, Alaizari N, et al. Aloe vera in treatment of oral submucous fibrosis: a systematic review and meta-analysis. J Oral Pathol Med. 2019;48(2):99-107.

98. Ali S, Wahbi W. The efficacy of aloe vera in management of oral lichen planus: a systematic review and meta-analysis. Oral Dis. 2017;23(7):913-8.

99. Dhingra K. Aloe vera herbal dentifrices for plaque and gingivitis control: a systematic review. Oral Dis. 2014;20(3):254-67.

100. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? PLoS medicine. 2010;7(9):e1000326.

101. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. Int J Evid Based Healthc. 2015;13(3):132-40.

102. Whiting P, Savović J, Higgins J, Caldwell D, Reeves B, Shea B, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol. 2018;69:225-34.

103. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017;358:j4008.

104. Pieper D, Puljak L, González-Lorenzo M, Minozzi S. Minor differences were found between AMSTAR 2 and ROBIS in the assessment of systematic reviews including both randomized and nonrandomized studies. J Clin Epidemiol. 2019;108:26-33.

105. Buehn S, Mathes T, Prengel P, Wegewitz U, Ostermann T, Robens S, et al. The risk of bias in systematic reviews tool showed fair reliability and good construct validity. J Clin Epidemiol. 2017;91:121-8.

106. Banzi R, Cinquini M, Gonzalez-Lorenzo M, Pecoraro V, Capobussi M, Minozzi S. Quality assessment versus risk of bias in systematic reviews: AMSTAR and ROBIS had similar reliability but differed in their construct and applicability. J Clin Epidemiol. 2018;99:24-32.

107. Giannakou K, Evangelou E, Yiallouros P, Christophi CA, Middleton N, Papatheodorou E, et al. Risk factors for gestational diabetes: An umbrella review of meta-analyses of observational studies. PLoS One. 2019;14(4):e0215372.

108. Veronese N, Solmi M, Caruso MG, Giannelli G, Osella AR, Evangelou E, et al. Dietary fiber and health outcomes: an umbrella review of systematic reviews and meta-analyses. Am J Clin Nutr. 2018;107(3):436-44.

109. Sullivan GM, Feinn R. Using Effect Size-or Why the P Value Is Not Enough. Journal of graduate medical education. 2012;4(3):279-82.

110. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-88.

111. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-60.

112. IntHout J, Ioannidis JP, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open. 2016;6(7):e010247.

113. Ioannidis JP, Trikalinos TA. An exploratory test for an excess of significant

findings. Clin Trials. 2007;4(3):245-53.

114. Ioannidis JP. Clarifications on the application and interpretation of the test for excess significance and its extensions. J Math Psychol. 2013;57(5):184-7.

115. Tsilidis KK, Papatheodorou SI, Evangelou E, Ioannidis JPA. Evaluation of excess statistical significance in meta-analyses of 98 biomarker associations with cancer risk. J Natl Cancer Inst. 2012;104(24):1867-78.

116. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-34.

117. Ioannidis JP, Trikalinos TA. The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. CMAJ. 2007;176(8):1091-6.

118. World Health Organization. The use of bedaquiline in the treatment of multidrug-resistant tuberculosis: interim policy guidance Geneva: World Health Organization; 2013 [cited 2019 20 October]. Available from: https://www.who.int/tb/publications/mdrtb-treatment-guideline/en/.

119. Fiocchi A, Pawankar R, Cuello-Garcia C, Ahn K, Al-Hammadi S, Agarwal A, et al. World allergy organization-mcmaster university guidelines for allergic disease prevention (GLAD-P): probiotics. World Allergy Organ J. 2015;8(1):4.

120. Dellinger RP, Levy MM, Rhodes A, Annane D, Gerlach H, Opal SM, et al. Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock, 2012. Intensive Care Med. 2013;39(2):165-228.

121. Agoritsas T, Merglen A, Heen AF, Kristiansen A, Neumann I, Brito JP, et al. Uptodate adherence to grade criteria for strong recommendations: an analytical survey. BMJ open. 2017;7(11):e018593-e.

122. De Vries FEE, Wallert ED, Solomkin JS, Allegranzi B, Egger M, Dellinger EP, et al. A systematic review and meta-analysis including GRADE qualification of the risk of surgical site infections after prophylactic negative pressure wound therapy compared with conventional dressings in clean and contaminated surgery. Medicine. 2016;95(36):e4673.

123. Poole R, Kennedy OJ, Roderick P, Fallowfield JA, Hayes PC, Parkes J. Coffee consumption and health: umbrella review of meta-analyses of multiple health outcomes. BMJ. 2017;359:j5024.

124. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, et al.

GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011;64(4):401-6.

125. Theodoratou E, Tzoulaki I, Zgaga L, Ioannidis JP. Vitamin D and multiple health outcomes: umbrella review of systematic reviews and meta-analyses of observational studies and randomised trials. BMJ. 2014;348:g2035.

126. Li X, Meng X, Timofeeva M, Tzoulaki I, Tsilidis KK, Ioannidis JP, et al. Serum uric acid levels and multiple health outcomes: umbrella review of evidence from observational studies, randomised controlled trials, and Mendelian randomisation studies. BMJ. 2017;357:j2376.

127. Rezende LFM, Sa TH, Markozannes G, Rey-Lopez JP, Lee IM, Tsilidis KK, et al. Physical activity and cancer: an umbrella review of the literature including 22 major anatomical sites and 770 000 cancer cases. Br J Sports Med. 2018;52(13):826-33.

128. Radua J, Ramella-Cravaro V, Ioannidis JP, Reichenberg A, Phiphopthatsanee N, Amir T, et al. What causes psychosis? An umbrella review of risk and protective factors. World Psychiatry. 2018;17(1):49-66.

129. Yu H, Zhong X, Gao P, Shi J, Wu Z, Guo Z, et al. The potential effect of metformin on cancer: An umbrella review. Front Endocrinol (Lausanne). 2019;10:617.

130. Neuenschwander M, Ballon A, Weber KS, Norat T, Aune D, Schwingshackl L, et al. Role of diet in type 2 diabetes incidence: umbrella review of meta-analyses of prospective observational studies. bmj. 2019;366:l2368.

131. Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. J R Stat Soc Ser A Stat Soc. 2009;172(1):137-59.

132. Al-Maweri SA, Nassani MZ, Alaizari N, Kalakonda B, Al-Shamiri HM, Alhajj MN, et al. Efficacy of aloe vera mouthwash versus chlorhexidine on plaque and gingivitis: a systematic review. Int J Dent Hyg. 2019;04:44-51.

133. Chen W, Zhang Y, Li X, Yang G, Liu JP. Chinese herbal medicine for diabetic peripheral neuropathy. Cochrane Database Syst Rev. 2013(10):CD007796.

134. Farahnik B, Sharma D, Alban J, Sivamani RK. Topical botanical agents for the treatment of psoriasis: a systematic review. Am J Clin Dermatol. 2017;18(4):451-68.

135. Farrugia CJ, Burke ES, Haley ME, Bedi KT, Gandhi A. The use of aloe vera in cancer radiation: an updated comprehensive review. Complement Ther Clin Pract.

2019;35:126-30.

136. Fifi AC, Axelrod CH, Chakraborty P, Saps M. Herbs and spices in the treatment of functional gastrointestinal disorders: a review of clinical trials. Nutrients. 2018;10(11):E1715.

137. Furness S, Worthington HV, Bryan G, Birchenough S, McMillan R. Interventions for the management of dry mouth: topical therapies. Cochrane Database Syst Rev. 2011;12:CD008934.

138. Gupta AK, Versteeg SG. Topical Treatment of Facial Seborrheic Dermatitis: A Systematic Review. Am J Clin Dermatol. 2017;18(2):193-213.

139. Hekmatpou D, Mehrabi F, Rahzani K, Aminiyan A. The effect of aloe vera clinical trials on prevention and healing of skin wound: a systematic review. Iran J Med Sci. 2019;44(1):1-9.

140. Marous MR, Flaten HK, Sledge B, Rietcheck H, Dellavalle RP, Suneja T, et al. Complementary and alternative methods for treatment of acne vulgaris: a systematic review. Curr Dermatol Rep. 2018;7(4):359-70.

141. Ramkumar D, Rao SSC. Efficacy and safety of traditional medical therapies for chronic constipation: systematic review. Am J Gastroenterol. 2005;100(4):936-71.

142. Gao Y, Jiang T, Mei S, Zhang S, Zhu C, Sun Y. Meta-analysis of aloe vera for the prevention and treatment of chemotherapy-induced phlebitis. Int J Clin Exp Med. 2016;9(6):9642-50.

143. Norman G, Christie J, Liu Z, Westby MJ, Jefferies JM, Hudson T, et al. Antiseptics for burns. Cochrane Database Syst Rev. 2017;7:CD011821.

144. Wang XT, Teng YJ, Ge XJ, Ma H, Yang XL, Chen WS, et al. Effectiveness of aloe vera for acute and chronic wounds: a systematic review. CJEBM. 2013;13(4):468-73.

145. Worthington HV, Clarkson JE, Bryan G, Furness S, Glenny AM, Littlewood A, et al. Interventions for preventing oral mucositis for patients with cancer receiving treatment. Cochrane Database Syst Rev. 2011(4):Cd000978.

146. Gorski LA. The 2016 infusion therapy standards of practice. Home Healthc Now. 2017;35(1):10-8.

147. Higginson R, Parry A. Phlebitis: treatment, care and prevention. Nursing times. 2011;107(36):18-21.

148. dos Reis PE, Silveira RC, Vasques CI, de Carvalho EC. Pharmacological interventions to treat phlebitis: systematic review. J Infus Nurs. 2009;32(2):74-9.

149. Berardi R, Piga A, Pulita F, Romagnoli E, Pietroselli D, Carle F, et al. Effective prevention of 5-fluorouracil-induced superficial phlebitis by ketoprofen lysine salt gel. Am J Med. 2003;115(5):415-7.

150. Thürlimann B, Bachmann I. Effective prevention of chemotherapy—induced phlebitis by low-dose heparin: a prospective randomised trial. Ann Oncol. 1992;3(4):311-3.

151. Ikeda S-i, Douchi T, Nagata Y. Use of heparin to lower the incidence of phlebitis induced by anti-neoplastic agents used in ovarian cancer. J Obstet Gynaecol Res. 2004;30(6):427-9.

152. Hamabe Y, Hanai A, Ishiguro H, Kuroda T, Hirota M, Nomura M, et al. Effects of steroid ointment application on chemotherapy-induced phlebitis: a randomized, double-blind, placebo-controlled clinical trial. Ann Oncol. 2017;28(suppl10):x155-65.

153. Bigdeli Shamloo MB, Nasiri M, Maneiy M, Dorchin M, Mojab F, Bahrami H, et al. Effects of topical sesame (Sesamum indicum) oil on the pain severity of chemotherapy-induced phlebitis in patients with colorectal cancer: a randomized controlled trial. Complement Ther Clin Pract. 2019;35:78-85.

154. Behnamfar N, Parsa Yekta Z, Mojab F, Kazem Naeini SM. The effect of nigella sativa oil on the prevention of phlebitis induced by chemotherapy: a clinical trial. Biomedicine (Taipei). 2019;9(3):20.

155. Zhang F, Jin W, Wang Y. The nursing research of frozen fresh potato topical treatment infusion leakage. Am J Nurs. 2018;7(4):143-6.

156. Mandal A, Raghu K. Study on incidence of phlebitis following the use of pheripheral intravenous catheter. J Family Med Prim Care. 2019;8(9):2827-31.

157. Sait M, Aguam A, Mohidin S, Al Eidraous S, Al Tabsh L, Anfinan N. Intravenous site complications for patients receiving chemotherapy: an observational study. Ann Short Reports.2(1032):1-4.

158. Tzolos E, Salawu A. Improving the frequency of visual infusion phlebitis (VIP) scoring on an oncology ward. BMJ Qual Improv Rep. 2014;3(1).

159. A World Health Organization resource. National List of Essential Medicines, 2018

– Thailand 2018 [cited 2020 30 January]. Available from: https://apps.who.int/medicinedocs/en/m/abstract/Js23739th/.

160. Sandercock P. The authors say: 'The data are not so robust because of heterogeneity' – so, How should I deal with this systematic review? Cerebrovasc Dis. 2011;31(6):615-20.

161. Davari P, Gorouhi F, Hsiao H, Fazel N. A comprehensive systematic review and meta-analysis of treatments for lichen planus. the 75th annual meeting of the society for investigative dermatology; May, 11-14; Scottsdale, AZ2016. p. S47-S.

162. Deng SQ, Brian M, Tony Z, Lu CJ, Xue CC. Herbal medicines in the topical management of psoriasis: a systematic review of clinical evidence. ASMR victorian student research symposium; Jun, 3; Melbourne, Australia2013. p. 226-7.

163. Mei S, Zhang S. Meta analysis of prevention and treatment of aloe for patients with chemotherapy induced phlebitis. Chin Nurs Res. 2016;30(10B):3627-31.

164. Shereen A, Maha A, Walaa W. Do nutraceuticals and herbal medicines have a role in managing oral lichen planus? A systematic review. J Oral Pathol Med. 2018;48:60.

165. Koch AK, Trifunovic-Koenig M, Klose P, Cramer H, Dobos G, Langhorst J. Herbal medicine in the treatment of irritable bowel syndrome-a systematic review. World Congress Integrative Medicine & Health; May, 3-5; Berlin, Germany2017.

166. Chang FY, Lu CL. Treatment of Irritable Bowel Syndrome Using Complementary and Alternative Medicine. Journal of the Chinese Medical Association. 2009;72(6):294-300.

167. Daniyal M, Akram M, Zainab R, Munir N, Shah SMA, Liu B, et al. Progress and prospects in the management of psoriasis and developments in phyto-therapeutic modalities. Dermatol Ther. 2019;32(3):e12866.

168. Esters P, Dignass A. Complementary therapies in inflammatory bowel diseases. Current Drug Targets. 2014;15(11):1079-88.

169. Grace OM, Simmonds MSJ, Smith GF, van Wyk AE. Therapeutic uses of Aloe L. (Asphodelaceae) in southern Africa. J Ethnopharmacol. 2008;119(3):604-14.

170. Keenan AV, Ferraiolo D. Insufficient evidence for effectiveness of any treatment for oral lichen planus. Evid Based Dent. 2011;12(3):85-6.

171. Subiksha PS. Various remedies for recurrent aphthous ulcer-a review. Journal of Pharmaceutical Sciences and Research. 2014;6(6):251-3.

172. Baccaglini L, Thongprasom K, Carrozzo M, Bigby M. Urban legends series: lichen planus. Oral Dis. 2013;19(2):128-43.

173. Feily A, Namazi MR. Aloe vera in dermatology: a brief review. G Ital Dermatol Venereol. 2009;144(1):85-91.

174. Jun S, Yoshimasa K, Takuya A, Noritaka O, Erika T, Emi Y, et al. Review of the current international consensus on burning mouth syndrome: Treatment options. J Oral Pathol Med. 2019;48(Suppl 1):19.

175. Ochmann K, Beer AM. Aloe vera gel: a literature research. Zeitschrift fur Phytotherapie. 2017;38(4):158-64.

176. Osso D, Kanani N. Antiseptic Mouth Rinses: An Update on Comparative Effectiveness, Risks and Recommendations. J Dent Hyg. 2013;87(1):10-8.

177. Pandey A, Singh S. Aloe vera: a systematic review of its industrial and ethno-medicinal efficacy. Int J Pharm Res Allied Sci. 2016;5(1):21-33.

178. Sidgwick GP, McGeorge D, Bayat A. A comprehensive evidence-based review on the role of topicals and dressings in the management of skin scarring. Archives of Dermatological Research. 2015;307(6):461-77.

179. Singab AN, El-Hefnawy HM, Esmat A, Gad HA, Nazeam JA. A systemic review on Aloe arborescens pharmacological profile: biological activities and pilot clinical trials. Phytother Res. 2015;29(12):1858-67.

180. Alebie G, Urga B, Worku A. Systematic review on traditional medicinal plants used for the treatment of malaria in Ethiopia: Trends and perspectives. Malar J. 2017;16 (1):307.

181. Bonchak JG, Thareja S, Chen SC, Quave CL. Botanical complementary and alternative medicine for pruritus: a systematic review. Curr Derm Rep. 2017;6(4):248-55.

182. Davis RH, Leitner MG, Russo JM. Aloe vera. a natural approach for treating wounds, edema and pain in diabetes. J Am Podiatr Med Assoc. 1988;78(2):60-8.

183. Ernst E, Pittler MH, Stevinson C. Complementary/alternative medicine in dermatology: Evidence-assessed efficacy of two diseases and two treatments. American Journal of Clinical Dermatology. 2002;3(5):341-8.

184. Heil K, Thomas R, Robertson G, Porter A, Milner R, Wood A. Freezing and non-freezing cold weather injuries: A systematic review. Br Med Bull. 2016;117(1):79-93.

185. Hon KL, Kung JSC, Ng WGG, Leung TF. Emollient treatment of atopic dermatitis: latest evidence and clinical considerations. Drugs in Context. 2018;7.

186. Jacobson LK, Johnson MB, Dedhia RD, Niknam-Bienia S, Wong AK. Impaired wound healing after radiation therapy: A systematic review of pathogenesis and treatment. JPRAS Open. 2017;13:92-105.

187. Lall N, Kishore N. Are plants used for skin care in South africa fully explored? J Ethnopharmacol. 2014;153(1):61-84.

188. Mollazadeh H, Mahdian D, Hosseinzadeh H. Medicinal plants in treatment of hypertriglyceridemia: a review based on their mechanisms and effectiveness. Phytomedicine. 2019;53:43-52.

189. Pazyar N, Yaghoobi R, Rafiee E, Mehrabian A, Feily A. Skin wound healing and phytomedicine: a review. Skin Pharmacol Physiol. 2014;27(6):303-10.

190. Rippon M, Perrin A, Darwood R, Ousey K. The potential benefits of using aloe vera in stoma patient skin care. Br J Nurs. 2017;26(5):S12-9.

191. Lee T, Dugoua JJ. Nutritional supplements and their effect on glucose control. Curr Diab Rep. 2011;11(2):142-8.

192. Shabanian S, Bahmani M, Asadi-Samani M. The medicinal plants effective on female hormones: a review of the native medicinal plants of Iran effective on estrogen, progesterone, and prolactin. JCPS. 2016;9(3):1270-6.

193. Amoo SO, Aremu AO, Van Staden J. Unraveling the medicinal potential of South African Aloe species. J Ethnopharmacol. 2014;153(1):19-41.

194. Arnold E, Clark CE, Lasserson TJ, Wu T. Herbal interventions for chronic asthma in adults and children. Cochrane Database Syst Rev. 2008(1):CD005989.

195. Bell PL, Gabriel V. Evidence based review for the treatment of post-burn pruritus. J Burn Care Res. 2009;30(1):55-61.

196. Chan ESY, Bautista DT, Zhu Y, You Y, Long JT, Li W, et al. Traditional Chinese herbal medicine for vascular dementia. Cochrane Database Syst Rev. 2018(12):CD010284.

197. Ganjalivand S, Haghshenas A, Lashkari F, Ghazanfarabadi M, Behzadi MR,

Chahkhoei M. Assessment of the variable types of burn dressings. International Journal of Pharmaceutical Research. 2018;10(1):264-9.

198. Gordon M, MacDonald JK, Parker CE, Akobeng AK, Thomas AG. Osmotic and stimulant laxatives for the management of childhood constipation. Cochrane Database Syst Rev. 2016(8):CD009118.

199. Heidarian E, Rouhi-Boroujeni H, Deris F, Rafieian-Kopaei M. Medicinal plants with multiple effects on cardiovascular diseases: a systematic review. Curr Pharm Des. 2017;23(7):999-1015.

200. Klotz T, Kurmis R, Munn Z, Heath K, Greenwood JE. The effectiveness of moisturizers in the management of burn scars following burn injury: a systematic review. JBI Database System Rev Implement Rep. 2015;13(10):291-315.

201. Rahmani S, Yazdi AP. Use of herbal medication in osteoarthritis: a systematic review. Avicenna J Phytomed. 2015;5:124.

202. Pandey A, Tripathi P, Pandey R, Srivatava R, Goswami S. Alternative therapies useful in the management of diabetes: a systematic review. J Pharm Bioallied Sci. 2011;3(4):504-12.

203. Norman G, Westby MJ, Rithalia AD, Stubbs N, Soares MO, Dumville JC. Dressings and topical agents for treating venous leg ulcers. Cochrane Database Syst Rev. 2018(6):CD012583.

204. Hollinger JC, Angra K, Halder RM. Are natural ingredients effective in the management of hyperpigmentation? a systematic review. J Clin Aesthet Dermatol. 2018;11(2):28-37.

205. Asaadi N, Kariman N, Shahrahmani H, Ghalandari S, Khodakarami N. A systematic review of clinical trials in the treatment of sore nipple and nipple pain in breastfeeding women. IJOGI. 2016;19(6):22-33.

206. As'adi N, Kariman N. Herbal prevention and treatment of nipple trauma and/or pain in Iranian studies: A systematic review. Journal of HerbMed Pharmacology. 2018;7(3):168-75.

207. Bahramsoltani R, Farzaei MH, Rahimi R. Medicinal plants and their natural components as future drugs for the treatment of burn wounds: an integrative review. Arch Dermatol Res. 2014;306(7):601-17.

208. Bindlish S, Shubrook JH, Jr. Dietary and botanical supplement therapy in diabetes. Osteopathic Family Physician. 2014;6(6):8-15.

209. Bolderston A, Lloyd NS, Wong RKS, Holden L, Robb-Blenderman L. The prevention and management of acute skin reactions related to radiation therapy: a systematic review and practice guideline. Support Care Cancer. 2006;14(8):802-17.

210. Bradley M, Cullum N, Nelson EA, Petticrew M, Sheldon T, Torgerson D. Systematic reviews of wound care management: (2) dressings and topical agents used in the healing of chronic wounds. Health Technology Assessment. 1999;3(17 II):iii-126.

211. Butcher K, Williamson K. Management of erythema and skin preservation; advice for patients receiving radical radiotherapy to the breast: a systematic literature review. J Radiother Pract. 2012;11(1):44-54.

212. Chan RJ, Webster J, Chung B, Marquart L, Ahmed M, Garantziotis S. Prevention and treatment of acute radiation-induced skin reactions: a systematic review and meta-analysis of randomized controlled trials. BMC Cancer. 2014;14(1):53.

213. Cheng S, Kirtschig G, Cooper S, Thornhill M, Leonardi-Bee J, Murphy R. Interventions for erosive lichen planus affecting mucosal sites. Cochrane Database Syst Rev. 2012;2:CD008092.

214. Deng S, May BH, Zhang AL, Lu C, Xue CC. Plant extracts for the topical management of psoriasis: a systematic review and meta-analysis. Br J Dermatol. 2013;169(4):769-82.

215. Ferreira E, Vasques C, Gadia R, Chan R, Guerra E, Mezzomo L, et al. Topical interventions to prevent acute radiation dermatitis in head and neck cancer patients: a systematic review. Support Care Cancer. 2017;25(3):1001-11.

216. Ghosh S, Gupta S. Interventions for the management of oral lichen planus: a review of the conventional and novel therapies. Oral Dis. 2017;23(8):1029-42.

217. Herman A, Herman AP. Topically used herbal products for the treatment of psoriasis - mechanism of action, drug delivery, clinical studies. Planta Med. 2016;82(17):1447-55.

218. Koukourakis GV, Kelekis N, Kouvaris J, Beli IK, Kouloulias VE. Therapeutics interventions with anti-inflammatory creams in post radiation acute skin reactions: a systematic review of most important clinical trials. Recent Pat Inflamm Allergy Drug

Discov. 2010;4(2):149-58.

219. Kumar S, Juresic E, Barton M, Shafiq J. Management of skin toxicity during radiation therapy: A review of the evidence. J Med Imaging Radiat Oncol. 2010;54(3):264-79.

220. Langhorst J. Complementary and alternative medicine treatments in inflammatory bowel diseases. [German]. J Gastroenterol Hepatol Erkr. 2016;14(1):9-16.

221. Liu JP, Yang M, Liu Y, Wei ML, Grimsgaard S. Herbal medicines for treatment of irritable bowel syndrome. Cochrane Database Syst Rev. 2006(1):CD004116.

222. Liu ZL, Liu JP, Zhang AL, Wu Q, Ruan Y, Lewith G, et al. Chinese herbal medicines for hypercholesterolemia. Cochrane Database Syst Rev. 2011(7):CD008305.

223. Lodi G, Carrozzo M, Furness S, Thongprasom K. Interventions for treating oral lichen planus: a systematic review. Br J Dermatol. 2012;166(5):938-47.

224. Magin PJ, Adams J, Pond CD, Smith W. Topical and oral CAM in acne: a review of the empirical evidence and a consideration of its context. Complement Ther Med. 2006;14(1):62-76.

225. Miroddi M, Navarra M, Calapai F, Mancari F, Giofre SV, Gangemi S, et al. Review of clinical pharmacology of Aloe vera L. in the treatment of psoriasis. Phytother Res. 2015(5):648-55.

226. Moore Z, Cowman S. A systematic review of wound cleansing for pressure ulcers. J Clin Nurs. 2008;17(15):1963-72.

227. Moore ZE, Cowman S. Wound cleansing for pressure ulcers. Cochrane Database Syst Rev. 2013(3):CD004983.

228. Muthusamy RC, Dharman S. Use of aloe vera in the treatment of oral lichen planus-a systematic review. Int J Pharm Bio Sci. 2016;7(1):146-52.

229. Ng SC, Lam YT, Tsoi KK, Chan FK, Sung JJ, Wu JC. Systematic review: the efficacy of herbal therapy in inflammatory bowel disease. Aliment Pharmacol Ther. 2013;38(8):854-63.

230. Niazi A, Rahimi VB, Soheili-Far S, Askari N, Rahmanian-Devin P, Sanei-Far Z, et al. A systematicreview on prevention and treatment of nipple pain and fissure: are they curable? J Pharmacopuncture. 2018;21(3):139-50.

231. Norman G, Dumville JC, Mohapatra DP, Owens GL, Crosbie EJ. Antibiotics and

antiseptics for surgical wounds healing by secondary intention. Cochrane Database Syst Rev. 2016(3):CD011712.

232. Rahimi R, Abdollahi M. Herbal medicines for the management of irritable bowel syndrome: a comprehensive review. World J Gastroenterol. 2012;18(7):589-600.

233. Rahimi R, Abdollahi M. A systematic review of the topical drugs for post hemorrhoidectomy pain. Int J Pharmacol. 2012;8(7):628-37.

234. Rahimi R, Nikfar S, Abdollahi M. Induction of clinical response and remission of inflammatory bowel disease by use of herbal medicines: a meta-analysis. World J Gastroenterol. 2013;19(34):5738-49.

235. Rashidi AA, Mirhashemi SM, Taghizadeh M, Sarkhail P. Iranian medicinal plants for diabetes mellitus: a systematic review. Pak J Biol Sci. 2013;16(9):401-11.

236. Richardson J, Smith JE, McIntyre M, Thomas R, Pilkington K. Aloe vera for preventing radiation-induced skin reactions: a systematic literature review. Clin Oncol (R Coll Radiol). 2005;17(6):478-84.

237. Shahrahmani H, Kariman N, Jannesari S, Ghalandari S, Asadi N. A systematic review on the type of treatment methods to reduce pain and improve wound healing in Iran. IJOGI. 2016;19(9):17-31.

238. Shi J, Hu H, Harnett J, Zheng X, Liang Z, Wang YT, et al. An evaluation of randomized controlled trials on nutraceuticals containing traditional Chinese medicines for diabetes management: A systematic review. Chinese Medicine (United Kingdom). 2019;14(1).

239. Smith N, Weymann A, Tausk FA, Gelfand JM. Complementary and alternative medicine for psoriasis: a qualitative review of the clinical trial literature. J Am Acad Dermatol. 2009;61(5):841-56.

240. Suresh SS, Chokshi K, Desai S, Malu R, Chokshi A. Medical management of oral lichen planus: a systematic review. J Clin Diagn Res. 2016;10(2):ZE10-ZE5.

241. Thongprasom K, Carrozzo M, Furness S, Lodi G. Interventions for treating oral lichen planus. Cochrane Database Syst Rev. 2011(7):CD001168.

242. Vermeulen H, Ubbink D, Goossens A, de Vos R, Legemate D. Dressings and topical agents for surgical wounds healing by secondary intention. Cochrane Database Syst Rev. 2004(2):Cd003554.

243. Vogler BK, Ernst E. Aloe vera: a systematic review of its clinical effectiveness. Br J Gen Pract. 1999;49(447):823-8.

244. Yarom N, Ariyawardana A, Hovan A, Barasch A, Jarvis V, Jensen SB, et al. Systematic review of natural agents for the management of oral mucositis in cancer patients. Supportive Care in Cancer. 2013;21(11):3209-21.

245. Yeh GY, Eisenberg DM, Kaptchuk TJ, Phillips RS. Systematic review of herbs and dietary supplements for glycemic control in diabetes. Diabetes Care. 2003;26(4):1277-94.

246. Zhang Y, Zhang S, Shao X. Topical agent therapy for prevention and treatment of radiodermatitis: a meta-analysis. Support Care Cancer. 2013;21(4):1025-31.

VITA

NAME	Saranrat Sadoyu
DATE OF BIRTH	19 March 1990
PLACE OF BIRTH	Nakhonratchasima
INSTITUTIONS ATTENDED	Srinakarinwirot university
HOME ADDRESS	Pakchong, Nakhonratchasima

Appendix

Appendix 1 Search strategy

Databases	Search strategy
AMED via Ovid Results = 212	1. aloe.mp. [mp=abstract, heading words, title] 2. aloe vera.mp. [mp=abstract, heading words, title] 3. 1 OR 2
CINAHL Plus via EBSCOhost Results = 33	1. aloe or aloe vera 2. systematic review or meta-analysis 3. 1 AND 2
Cochrane database of Systematic reviews Results = 37	1. aloe OR aloe vera 2. systematic review OR meta analysis 3. 1 AND 2 in Cochrane Database of Systematic Reviews
Embase via Ovid Results = 157	1. aloe.mp. or Aloe vera extract/ or aloe emodin/ or Aloe vera/ or Aloe barbadensis extract/ or Aloe/ or aloe emodin anthrone/ or aloe vera.mp. 2. systematic review or meta-analysis 3. 1 AND 2
PubMed Results = 70	1. aloe OR aloe vera 2. systematic review OR meta analysis 3. 1 AND 2 ("aloe"[MeSH Terms] OR "aloe"[All Fields]) AND (("systematic review"[Publication Type] OR "systematic reviews as topic"[MeSH Terms] OR "systematic review"[All Fields]) OR ("meta-analysis"[Publication Type] OR "meta-analysis as topic"[MeSH Terms] OR "meta-analysis"[All Fields]))
Scopus Results = 182	1. TITLE-ABS-KEY (“Aloe” OR “Aloe vera”) 2. TITLE-ABS-KEY (“systematic review” OR “meta-analysis") 3. 1 AND 2

Appendix 2 AMSTAR 2 checklist

1. Did the research questions and inclusion criteria for the review include the components of PICO?

For Yes:

- ☐ Population
- ☐ Intervention
- ☐ Comparator group
- ☐ Outcome

Optional (recommended)

- ☐ Timeframe for follow-up

- ☐ Yes
- ☐ No

2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

For Partial Yes:

The authors state that they had a written protocol or guide that included ALL the following:

- ☐ review question (s)
- ☐ a search strategy
- ☐ inclusion/exclusion criteria
- ☐ a risk of bias assessment

For Yes:

As for partial yes, plus the protocol should be registered and should also have specified:

- ☐ a meta-analysis/synthesis plan, if appropriate, *and*
- ☐ a plan for investigating causes of heterogeneity
- ☐ justification for any deviations from the protocol

- ☐ Yes
- ☐ Partial Yes
- ☐ No

3. Did the review authors explain their selection of the study designs for inclusion in the review?

For Yes, the review should satisfy ONE of the following:

- ☐ Explanation for including only RCTs
- ☐ OR Explanation for including only NRSI
- ☐ OR Explanation for including both RCTs and NRSI

- ☐ Yes
- ☐ No

4. Did the review authors use a comprehensive literature search strategy?

For Partial Yes (all the following):

- ☐ searched at least 2 databases (relevant to research question)
- ☐ provided key word and/or search strategy
- ☐ justified publication restrictions (e.g. language)

For Yes, should also have (all the following)

- ☐ searched the reference lists / bibliographies of included studies
- ☐ searched trial/study registries
- ☐ included/consulted content
- ☐ where relevant, searched for grey literature
- ☐ conducted search within 24 months of completion of the review

- ☐ Yes
- ☐ Partial yes
- ☐ No

5. Did the review authors perform study selection in duplicate?

For Yes, either ONE of the following:

- ☐ at least two reviewers independently agreed on selection of eligible studies and achieved consensus on which studies to include
- ☐ OR two reviewers selected a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder selected by one reviewer

☐ Yes
☐ No

6. Did the review authors perform data extraction in duplicate?

For Yes, either ONE of the following:

- ☐ at least two reviewers achieved consensus on which data to extract from included studies
- ☐ OR two reviewers extracted data from a sample of eligible studies and achieved good agreement (at least 80 percent), with the remainder extracted by one reviewer.

☐ Yes
☐ No

7. Did the review authors provide a list of excluded studies and justify the exclusions?

For Partial Yes:

- ☐ provided a list of all potentially relevant studies that were read in full-text form but excluded from the review

For Yes, must also have:

- ☐ Justified the exclusion from, the review of each potentially relevant study

☐ Yes
☐ Partial yes
☐ No

8. Did the review authors describe the included studies in adequate detail?

For Partial Yes (all the following):

- ☐ described populations
- ☐ described interventions
- ☐ described comparators
- ☐ described outcomes
- ☐ described research designs

For Yes, should also have all the following

- ☐ described population in detail
- ☐ described interventions in detail (including doses where relevant)
- ☐ described comparator in detail (including doses where relevant)
- ☐ described study's setting
- ☐ timeframe for follow-up

☐ Yes
☐ Partial yes
☐ No

9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

RCTs

For Partial Yes, must have assessed RoB from:

- ☐ unconcealed allocation, and
- ☐ lack of blinding of patients and assessors when assessing outcomes (unnecessary for objective outcomes such as all-cause mortality)

For Yes, must also have assessed RoB from:

- ☐ allocation sequence that was not truly random, and
- ☐ selection of the reported result from among multiple measurement or analyses of a specified outcome

- ☐ Yes
- ☐ Partial Yes
- ☐ No

NRSI

For Partial Yes, must have assessed RoB

- ☐ from confounding, and
- ☐ from selection bias

For Yes, must also have assessed RoB

- ☐ methods used to ascertain exposures and outcomes, and
- ☐ selection of the reported result from among multiple measurement or analyses of a specified outcome

- ☐ Yes
- ☐ Partial Yes
- ☐ No
- ☐ Includes only RCTs

10. Did the review authors report on the sources of funding for the studies included in the review?

For Yes

- ☐ Must have reported on the sources of funding for individual studies included in the review. Note: Reporting that the reviewers looked for this information but it was not reported by study authors also qualifies

- ☐ Yes
- ☐ No

11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?

RCTs

For Yes:

- ☐ The authors justified combining the data in a meta-analysis
- ☐ AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present.
- ☐ AND investigated the causes of any heterogeneity

- ☐ Yes
- ☐ No
- ☐ No meta- analysis conducted

For NRSI

For Yes:

- ☐ The authors justified combining the data in a meta-analysis
- ☐ AND they used an appropriate weighted technique to combine study results and adjusted for heterogeneity if present.
- ☐ AND they statistically combined effect estimates from NRSI that were adjusted for confounding, rather than combining raw data, or justified combining raw data when adjusted effect estimates were not available.
- ☐ AND reported separate summary estimates for RCTs and NRSI separately when both were included in the review.

☐ Yes ☐ No ☐ No meta-analysis conducted

12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?

For Yes

- ☐ included only low risk of bias RCTs
- ☐ OR, if the pooled estimate was based on RCTs and/or NRSI at variable RoB, the authors performed analyses to investigate possible impact of RoB on summary estimates of effect.

☐ Yes ☐ No ☐ No meta-analysis conducted

13. Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?

For Yes

- ☐ included only low risk of bias RCTs
- ☐ OR, if RCTs with moderate or high RoB, or NRSI were included the review provided a discussion of the likely impact of RoB on the results

☐ Yes ☐ No

14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

For Yes

- ☐ There was no significant heterogeneity in the results
- ☐ OR if heterogeneity was present the authors performed an investigation of sources of any heterogeneity in the results and discussed the impact of this on the results of the review

☐ Yes ☐ No

15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?

For Yes

- ☐ performed graphical or statistical tests for publication bias and discussed the likelihood and magnitude of impact of publication bias

☐ Yes
☐ No
☐ No meta-analysis conducted

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

- ☐ The authors reported no competing interests OR
- ☐ The authors described their funding sources and how they managed potential conflicts of interest

☐ Yes
☐ No

Appendix 3 Studies excluded after full-text revision, with reasons for exclusion

Author, Year	Title	Reason for exclusion
Davari, 2012(161)	A comprehensive systematic review and meta-analysis of treatments for lichen planus	full-text could not be retrieved
Deng, 2013(162)	Herbal medicines in the topical management of psoriasis: A systematic review of clinical evidence	full-text could not be retrieved
Mei, 2016(163)	Meta analysis of prevention and treatment of Aloe for patients with chemotherapy induced phlebitis	full-text could not be retrieved
Shereen, 2019(164)	Do nutraceuticals and herbal medicines have a role in managing oral lichen planus? A systematic review	full-text could not be retrieved
Koch, 2017(165)	Herbal medicine in the treatment of irritable bowel syndrome-a systematic review	full-text could not be retrieved
Chang, 2009(166)	Treatment of Irritable Bowel Syndrome Using Complementary and Alternative Medicine	not a systematic review/meta-analysis
Daniyal, 2019(167)	Progress and prospects in the management of psoriasis and developments in phyto-therapeutic modalities	not a systematic review/meta-analysis
Esters, 2014(168)	Complementary therapies in inflammatory bowel diseases	not a systematic review/meta-analysis
Grace, 2008(169)	Therapeutic uses of Aloe L. (Asphodelaceae) in southern Africa	not a systematic review/meta-analysis
Keenan, 2011(170)	Insufficient evidence for effectiveness of any treatment for oral lichen planus	not a systematic review/meta-analysis
Subiksha, 2014(171)	Various remedies for recurrent aphthous ulcer-a review	not a systematic review/meta-analysis
Baccaglini, 2013(172)	Urban legends series: Lichen planus	systematic review based on other systematic reviews/meta-analyses
Feily, 2009(173)	Aloe vera in dermatology: A brief review	systematic review based on other systematic reviews/meta-analyses
Jun, 2019(174)	Review of the current international consensus on burning mouth syndrome: Treatment options	systematic review based on other systematic

Author, Year	Title	Reason for exclusion
		reviews/meta-analyses
Ochmann, 2017(175)	Aloe vera gel: A literature research	systematic review based on other systematic reviews/meta-analyses
Osso, 2013(176)	Antiseptic Mouth Rinses: An Update on Comparative Effectiveness, Risks and Recommendations	systematic review based on other systematic reviews/meta-analyses
Pandey, 2016(177)	Aloe Vera: A systematic review of its Industrial and ethno-medicinal efficacy	systematic review based on other systematic reviews/meta-analyses
Radha, 2015(26)	Evaluation of biological properties and clinical effectiveness of Aloe vera: A systematic review	systematic review based on other systematic reviews/meta-analyses
Sidgwick, 2015(178)	A comprehensive evidence-based review on the role of topicals and dressings in the management of skin scarring	systematic review based on other systematic reviews/meta-analyses
Singab, 2015(179)	A systemic review on Aloe arborescens pharmacological profile: biological activities and pilot clinical trials	systematic review based on other systematic reviews/meta-analyses
Ulbricht, 2007(51)	An evidence-based systematic review of Aloe vera by the natural standard research collaboration	systematic review based on other systematic reviews/meta-analyses
Alebie, 2017(180)	Systematic review on traditional medicinal plants used for the treatment of malaria in Ethiopia: Trends and perspectives	not a systematic review/meta-analysis of clinical trials
Bonchak, 2017(181)	Botanical complementary and alternative medicine for pruritus: a systematic review	not a systematic review/meta-analysis of clinical trials
Davis, 1988(182)	Aloe vera. A natural approach for treating wounds, edema and pain in diabetes	not a systematic review/meta-analysis of clinical trials
Ernst, 2002(183)	Complementary/alternative medicine in dermatology: Evidence-assessed efficacy of two diseases and two	not a systematic review/meta-analysis of

Author, Year	Title	Reason for exclusion
	treatments	clinical trials
Heil, 2016(184)	Freezing and non-freezing cold weather injuries: a systematic review	not a systematic review/meta-analysis of clinical trials
Hon, 2018(185)	Emollient treatment of atopic dermatitis: latest evidence and clinical considerations	not a systematic review/meta-analysis of clinical trials
Jacobson, 2017(186)	Impaired wound healing after radiation therapy: A systematic review of pathogenesis and treatment	not a systematic review/meta-analysis of clinical trials
Lall, 2014(187)	Are plants used for skin care in South Africa fully explored?	not a systematic review/meta-analysis of clinical trials
Mollaza deh, 2019(188)	Medicinal plants in treatment of hypertriglyceridemia: A review based on their mechanisms and effectiveness	not a systematic review/meta-analysis of clinical trials
Pazyar, 2014(189)	Skin wound healing and phytomedicine: A review	not a systematic review/meta-analysis of clinical trials
Rippon, 2017(190)	The potential benefits of using aloe vera in stoma patient skin care	not a systematic review/meta-analysis of clinical trials
Lee, 2011(191)	Nutritional supplements and their effect on glucose control	not a systematic review/meta-analysis of clinical trials
Shabanian, 2016(192)	The medicinal plants effective on female hormones: A review of the native medicinal plants of Iran effective on estrogen, progesterone, and prolactin	not a systematic review/meta-analysis of clinical trials
Amoo, 2014(193)	Unraveling the medicinal potential of South African Aloe species	not report treatment effects of *Aloe vera*
Arnold, 2008(194)	Herbal interventions for chronic asthma in adults and children	not report treatment effects of *Aloe vera*
Bell,	Evidence based review for the treatment of post-burn	not report treatment

Author, Year	Title	Reason for exclusion
2009(195)	pruritus	effects of *Aloe vera*
Chan, 2018(196)	Traditional Chinese herbal medicine for vascular dementia	not report treatment effects of *Aloe vera*
Ganjalivand, 2018(197)	Assessment of the variable types of burn dressings	not report treatment effects of *Aloe vera*
Gordon, 2016(198)	Osmotic and stimulant laxatives for the management of childhood constipation	not report treatment effects of *Aloe vera*
Rouhi-Boroujeni, 2017(199)	Medicinal plants with multiple effects on cardiovascular diseases: A systematic review	not report treatment effects of *Aloe vera*
Klotz, 2015(200)	The effectiveness of moisturizers in the management of burn scars following burn injury: A systematic review	not report treatment effects of *Aloe vera*
Rahmani, 2015(201)	Use of herbal medication in osteoarthritis: A systematic review	not report treatment effects of *Aloe vera*
Pandey, 2011(202)	Alternative therapies useful in the management of diabetes: A systematic review	not report treatment effects of *Aloe vera*
Prasad, 2017	Management of chronic constipation in patients with diabetes mellitus	not report treatment effects of *Aloe vera*
Norman, 2018(203)	Dressings and topical agents for treating venous leg ulcers	not report treatment effects of *Aloe vera*
Hollinger, 2018(204)	Are natural ingredients effective in the management of hyperpigmentation? A systematic review	not the largest systematic review/ meta-analysis
Asaadi, 2016(205)	A systematic review of clinical trials in the treatment of sore nipple and nipple pain in breastfeeding women.	not the largest systematic review/ meta-analysis
As'adi, 2018(206)	Herbal prevention and treatment of nipple trauma and/or pain in Iranian studies: A systematic review	not the largest systematic review/meta-analysis
Bahramsoltani,	Medicinal plants and their natural components as future drugs for the treatment of burn wounds: An	not the largest systematic review/

Author, Year	Title	Reason for exclusion
2014(207)	integrative review	meta-analysis
Bindlish, 2014(208)	Dietary and botanical supplement therapy in diabetes	not the largest systematic review/meta-analysis
Bolderston, 2006(209)	The prevention and management of acute skin reactions related to radiation therapy: A systematic review and practice guideline	not the largest systematic review/ meta-analysis
Bradley, 1999(210)	Systematic reviews of wound care management: (2) dressings and topical agents used in the healing of chronic wounds	not the largest systematic review/meta-analysis
Burusapat, 2018(40)	Topical Aloe Vera Gel for Accelerated Wound Healing of Split-Thickness Skin Graft Donor Sites: A Double-Blind, Randomized, Controlled Trial and Systematic Review	not the largest systematic review/ meta-analysis
Butcher, 2012(211)	Management of erythema and skin preservation; advice for patients receiving radical radiotherapy to the breast: A systematic literature review	not the largest systematic review/ meta-analysis
Chan, 2014(212)	Prevention and treatment of acute radiation-induced skin reactions: A systematic review and meta-analysis of randomized controlled trials	not the largest systematic review/ meta-analysis
Cheng, 2012(213)	Interventions for erosive lichen planus affecting mucosal sites	not the largest systematic review/ meta-analysis
Dat, 2012(71)	Aloe vera for treating acute and chronic wounds	not the largest systematic review/ meta-analysis
Deng, 2013(214)	Plant extracts for the topical management of psoriasis: A systematic review and meta-analysis	not the largest systematic review/ meta-analysis
Dhingra, 2014(99)	Aloe vera herbal dentifrices for plaque and gingivitis control: a systematic review	not the largest systematic review/ meta-analysis
Dick,	Reduction of Fasting Blood Glucose and Hemoglobin	not the largest

Author, Year	Title	Reason for exclusion
2016(92)	A1c Using Oral Aloe Vera: A Meta-Analysis	systematic review/ meta-analysis
Ferreira, 2017(215)	Topical interventions to prevent acute radiation dermatitis in head and neck cancer patients: a systematic review	not the largest systematic review/ meta-analysis
Ghosh, 2017(216)	Interventions for the management of oral lichen planus: a review of the conventional and novel therapies	not the largest systematic review/ meta-analysis
Herman, 2016(217)	Topically used herbal products for the treatment of psoriasis - Mechanism of action, drug delivery, clinical studies	not the largest systematic review/ meta-analysis
Koukou rakis, 2010(218)	Therapeutics interventions with anti-inflammatory creams in post radiation acute skin reactions: A systematic review of most important clinical trials	not the largest systematic review/ meta-analysis
Kumar, 2010(219)	Management of skin toxicity during radiation therapy: A review of the evidence	not the largest systematic review/ meta-analysis
Langhorst, 2016(220)	Complementary and alternative medicine treatments in inflammatory bowel diseases.	not the largest systematic review/ meta-analysis
Langhorst, 2015(94)	Systematic review of complementary and alternative medicine treatments in inflammatory bowel diseases	not the largest systematic review/ meta-analysis
Liu, 2006(221)	Herbal medicines for treatment of irritable bowel syndrome	not the largest systematic review/ meta-analysis
Liu, 2011(222)	Chinese herbal medicines for hypercholesterolemia	not the largest systematic review/ meta-analysis
Lodi, 2012(223)	Interventions for treating oral lichen planus: A systematic review	not the largest systematic review/ meta-analysis
Maenthai	The efficacy of aloe vera used for burn wound	not the largest

Author, Year	Title	Reason for exclusion
song, 2007(10)	healing: a systematic review	systematic review/ meta-analysis
Magin, 2006(224)	Topical and oral CAM in acne: a review of the empirical evidence and a consideration of its context	not the largest systematic review/ meta-analysis
Miroddi, 2015(225)	Review of Clinical Pharmacology of Aloe vera L. in the Treatment of Psoriasis	not the largest systematic review/ meta-analysis
Moore, 2008(226)	A systematic review of wound cleansing for pressure ulcers	not the largest systematic review/ meta-analysis
Moore, 2013(227)	Wound cleansing for pressure ulcers	not the largest systematic review/ meta-analysis
Muthu samy, 2016(228)	Use of aloe vera in the treatment of oral lichen planus-a systematic review	not the largest systematic review/ meta-analysis
Nair, 2016(95)	Clinical effectiveness of aloe vera in the management of oral mucosal diseases-a systematic review	not the largest systematic review/ meta-analysis
Ng, 2013(229)	Systematic review: The efficacy of herbal therapy in inflammatory bowel disease	not the largest systematic review/ meta-analysis
Niazi, 2018(230)	A Systematic Review on Prevention and Treatment of Nipple Pain and Fissure: Are They Curable?	not the largest systematic review/ meta-analysis
Norman, 2016(231)	Antibiotics and antiseptics for surgical wounds healing by secondary intention	not the largest systematic review/ meta-analysis
Rahimi, 2012(232)	Herbal medicines for the management of irritable bowel syndrome: A comprehensive review	not the largest systematic review/ meta-analysis
Rahimi,	A systematic review of the topical drugs for post	not the largest

Author, Year	Title	Reason for exclusion
2012(233)	hemorrhoidectomy pain	systematic review/ meta-analysis
Rahimi, 2013(234)	Induction of clinical response and remission of inflammatory bowel disease by use of herbal medicines: A meta-analysis	not the largest systematic review/ meta-analysis
Rashidi, 2013(235)	Iranian medicinal plants for diabetes mellitus: a systematic review	not the largest systematic review/ meta-analysis
Richardson, 2005(236)	Aloe vera for preventing radiation-induced skin reactions: A systematic literature review	not the largest systematic review/ meta-analysis
Shahrah mani, 2016(237)	A systematic review on the type of treatment methods to reduce pain and improve wound healing in Iran.	not the largest systematic review/ meta-analysis
Shi, 2019(238)	An evaluation of randomized controlled trials on nutraceuticals containing traditional Chinese medicines for diabetes management: A systematic review	not the largest systematic review/meta-analysis
Smith, 2009(239)	Complementary and alternative medicine for psoriasis: A qualitative review of the clinical trial literature	not the largest systematic review/ meta-analysis
Suresh, 2016(240)	Medical management of oral lichen planus: A systematic review	not the largest systematic review/ meta-analysis
Thongpras om, 2011(241)	Interventions for treating oral lichen planus	not the largest systematic review/ meta-analysis
Vermeu len, 2005(242)	Dressings and topical agents for surgical wounds healing by secondary intention	not the largest systematic review/ meta-analysis
Vogler, 1999(243)	Aloe vera: a systematic review of its clinical effectiveness	not the largest systematic review/ meta-analysis

Author, Year	Title	Reason for exclusion
Yarom, 2013(244)	Systematic review of natural agents for the management of oral mucositis in cancer patients	not the largest systematic review/ meta-analysis
Yeh, 2003(245)	Systematic review of herbs and dietary supplements for glycemic control in diabetes	not the largest systematic review/ meta-analysis
Zhang, 2013(246)	Topical agent therapy for prevention and treatment of radiodermatitis: A meta-analysis	not the largest systematic review/ meta-analysis

Appendix 4 Characteristics and main findings of included systematic reviews

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
Dentistry							
Al-Maweri, 2019 (132)	gingivitis	358 healthy participants aged 18 years and older	mouthwash (99 and 100%) 4-30 days	chlorhexidine mouthwash (0.2 and 0.12%)	●plaque index ●gingival index ●gingival bleeding index	6 (RCTs)	●4 studies reported AV as effective as chlorhexidine in reducing plaque index, with no statistically significant differences between the two groups. However, 2 studies found chlorhexidine significantly more effective than AV. ●3 studies found AV and chlorhexidine equally efficient in reducing gingival inflammation, with no significant differences. However, 1 study found chlorhexidine slightly more effective than AV. ●Only 1 study reported slightly more

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
							effective of chlorhexidine in reducing the mean bleeding index than AV, with no statistically significant differences. ● In summary, AV shows promising results in reducing plaque and gingivitis scores.
Furness, 2011 (137)	dry mouth syndrome (xero -stomia)	123 patients with xerostomia due to radiotherapy	AV gel 1 week	CMC spray,canola oil spray, and mucin spray	● dryness of mouth ● patient preference	1 (RCT)	There were no statistically significant differences between any of the treatments about either oral dryness or patient preference.

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
Anti-diabetes							
Chen, 2013 (133)	diabetic peripheral neuropathy	40 patients with diabetic peripheral neuropathy	oral AV (not specified type of product) 3 months	no treatment	● improvement of symptom (global symptom score) ● change of motor and sensory nerve conduction velocity	1 (RCT)	AV showed a significantly better effect on peroneal motor nerve, peroneal nerve, and median sensory nerve conduction velocity. However, AV did not show a favorable effect on global symptom score improvement (RR 1.67, 95% CI 0.96 to 2.88).
Gastrointestinal (GI) disorders							
Fifi, 2018 (136)	Irritable bowel syndrome (IBS)	110 patients with IBS	oral AV (juice) 5 months	placebo	Patient quality of life (QOL)	1 (RCT)	There was no significance difference between the placebo and AV in improving QOL in patients with IBS.

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
Fifi, 2018 (136)	Gastro-esophageal reflux disease (GERD)	79 patients with GERD	oral AV (syrup) 4 weeks	omeprazole/ ranitidine	GERD symptoms (modified Reflux Disease Questionnaire)	1 (RCT)	The effect of AV on GERD symptoms (i.e. frequency of heartburn, food regurgitation, flatulence, belching, dysphagia) was comparable to that of ranitidine and omeprazole in relation to most symptoms.
Ramkumar, 2005 (141)	chronic consti-pation	35 patients with chronic constipation	oral AV combined with Psyllium and Celandin 28 days	placebo	laxative efficacy (stool consistency, bowel movement, abdominal pain)	1 (RCT)	The combination of AV, psyllium, and celandin was superior to placebo in the treatment of constipation. It showed more frequent bowel movement, softer stool with statically significant compared with placebo; however, abdominal pain was not reduced in either group.
Radiation-induced reactions							
Farrugi, 2019 (135)	radiation-induced skin	759 patients who have undergone	topical AV (cream, gel, ointment,	placebo, no treatment	●severity ●clinical	7 (controlled trials)	●In breast cancer patients, 4 studies suggested that AV was not found

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
	reactions	radiation therapy for the treatment of cancer	lotion) Each day following radiation treatment or 2-4 weeks after radiation		presentation		to be consistently effective for radiation adverse effects, but 1 study showed its effective for treatment of acute radiation proctitis. ● In head/neck, chest, and abdomen cancer patients, 2 studies concluded that AV showed protective properties against radiation-induced dermatitis, especially with cumulative radiation doses over 2,700 cGy which reported as statistically significant in 1 study.

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
Skin condition							
Farahnik, 2017 (134)	psoriasis	195 patients with plaque psoriasis	topical AV (cream, gel, extract, and home-care pack) 4-8 weeks	placebo and other treatment (0.1% TA cream)	● clearing of psoriatic plaques ● severity (erythema, induration, scaling, psoriasis area and severity index; PASI score) ● dermatology life quality index (DLQI)	4 (controlled trials)	● AV was significantly superior to control group in reduction of erythema, infiltration, and PASI score. ● 1 study reported that AV cream resulted in a significant clearing of psoriatic plaques higher than placebo (82.8 vs 7.7%, p<0.001). ● AV reduced desquamation significantly in 2 studies, but the difference was not significant between AV and placebo group in 1 study. ● 1 study reported that the mean DLQI scores decreased in both groups, but in comparison to

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
							baseline there was no significant difference between the two groups after 8 weeks.
Marous, 2018 (140)	acne vulgaris	84 patients with mild-severe acne vulgaris	topical AV (AV gel or combined with 2% Ocimum gratissimum gel) 4 weeks	placebo and 1% clindamycin gel	lesion reduction	1 (RCT)	The result showed that effect of combination of 25% AV with Ocimum gratissimum in lesion reduction was similar to 1% clindamycin, while the preparations containing 50 or 100% AV gel exhibited significantly better effects than the clindamycin.
Gupta, 2017 (138)	facial seborrheic dermatitis (SD)	46 patients with seborrheic dermatitis	topical AV (30% AV extract in emulsion) 4-6 weeks	placebo (aquosum cream)	● patient response ● signs and symptoms of SD (i.e. scaling, pruritus)	1 (RCT)	● AV was statistical significantly increased patient response (complete clearance and substantial improvement) assessed with a global scale and decreased symptoms (pruritus), sign (scaling), and number of facial sites.

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
							● AV are moderately recommended for use in the treatment of facial SD.
Wound healing							
Hekmatopou, 2019 (139)	cracked nipples	210 lactating women with breast fissure	topical AV (gel) 7-14 days	placebo or other treatments	● wound healing ● pain and discharge reduction	2 clinical trials	● AV was more effective than control group (e.g. breast milk, lanolin ointment) in healing cracked nipples. ● 1 study reported that the pain and damage of the nipple and discharge in AV group were much less than the control group which used lanolin ointment.

Abbreviations: AV – *Aloe vera*, N – number, CMC – carboxymethyl cellulose, RCT – randomized controlled trial, q-RCT – quasi-randomized controlled trial, CMC – carboxymethyl cellulose, TA – triamcinolone acetonide, RR – relative risk or risk ratio, MD – mean difference, FPG – fasting plasma glucose, HbA1c – hemoglobin A1c, IV – intravenous, GERD – Gastroesophageal reflux disease, GSRS – Gastrointestinal Symptom Rating Scale, IBSQOL – Irritable bowel syndrome Quality of Life, EuroQol – European Quality of Life Scale Community

Appendix 5 Characteristics and main findings of included meta-analyses

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
Dentistry							
Ali, 2017 (98)	oral lichen planus (OLP)	121 patients with symptomatic OLP	topical (gel, ointment and mouthwash) 4 and 8 weeks	placebo or corticosteroid (TA gel and paste)	● pain alleviation ● clinical improvement ● treatment response ● size of the lesion ● hospital anxiety-depression (HAD)	5 (4 RCTs, 1 q-RCT)	● The meta-analysis showed that AV is inferior to the control in pain alleviation and clinical improvement with statistically significant difference ● 2 studies showed that AV is superior to the control in treatment response. ● Only 1 study reported size of the lesion. They found that the size decreased significantly after treatment and after 2 months of discontinuation of the treatment. ● 1 study reported no changes of HAD Scale in both groups during the course of the study.

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
							● In summary, AV was effective in managing OLP in the AV group, not inferior when compared to placebo group and comparable to TA.
Al-Maweri, 2019 (97)	oral submucous fibrosis (OSF)	413 patients diagnosed clinically and/or histopatho-logically	5 studies used AV gel alone, 1 study used both topical and systemic AV 3 months	placebo or any medical interventions (1 study: combination of cortico-steroids, antioxidants, and hyaluronidase)	● objective: interincisal mouth opening, tongue protrusion, and cheek flexibility. ● subjective: pain/burning sensation	6 (RCTs)	● The results of meta-analysis showed statistically significant differences between AV and control groups in reducing pain/burning sensation at the end of the 1st and 2nd month, in favor of AV, but no significant differences were found at the end of the 3rd month. ● With regard to objective clinical outcomes, no statistically significant differences were found between the groups.

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
Anti-diabetes							
Suksom boon, 2016 (93)	glucose lowering	470 participants with prediabetes or early, non-treated diabetes and diagnosed type 2 diabetes	oral AV (raw crushed of AV leaves, juice, gel powder, extract capsules) 2-3 months	placebo or no treatment	● glycemic control (FPG, HbA1C)	8 (RCTs)	● The meta-analysis showed that in prediabetes, AV significantly lowered FPG only, with no effect on HbA1c. ● In type 2 diabetes, AV showed significant improvement in HbA1c, but only a marginal in lowered FPG (p=0.05). ● In summary, AV showed a possible effect on glycemic control in prediabetes and type 2 diabetes.
Zhang, 2016(12)	glucose and lipid lowering	415 participants with pre-diabetes and early untreated diabetes	oral AV (juice, powder, capsules) 6-12 weeks	placebo	● FPG, HbA1C, insulin level ● lipid profile (TC, TG, LDL, HDL)	5 (RCTs)	● The meta-analysis showed that AV significantly reduced the levels of FPG, HbA1c, TC, TG, LDL, and significantly increased HDL.

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
Gastrointestinal disorders							
Hong, 2018 (11)	irritable bowel syndrome (IBS)	151 patients with IBS	oral AV (juice, extract) 1-5 months	placebo	● Symptoms severity ● Response rate	3 (RCTs)	The meta-analysis showed that AV was significantly higher in IBS symptoms scores and response rate improvement compared to the placebo.
Phlebitis							
Gao, 2016 (142)	chemotherapy-induced phlebitis	4530 patients who receive IV chemotherapy with the grade of phlebitis ranged from 1st to 3rd degree	topical AV fresh leaves, juice, and gel; the AV had not been chemically treated (not reported duration of therapy)	conventional treatment or 50% $MgSO_4$	● preventive effect: incidence of phlebitis ● treatment effect: treatment efficiency (efficacy rate, cure rate)	16 clinical trials (8 RCTs, 4 q- RCTs, 4 unclear study design)	The meta-analysis showed that AV has some potential value for the prevention of chemotherapy-induced phlebitis. Overall incidence of CIP was lower in AV than in control group but no statistically significance was found; however, AV significantly reduced the occurrence of 2nd and 3rd-degree CIP, and improved total efficacy rate and the cure rate more than control group.

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
Zheng, 2014(14)	intravenous infusion-induced phlebitis	7465 patients who suffered from phlebitis of a peripheral limb vein - associated with the IV access device	topical AV alone or plus non-AV interventions 1-15 days	no treatment or the same non-AV interventions	● preventive effect: incidence of phlebitis ● treatment effect: rate of resolution of phlebitis	43 (35 RCTs, 8 q-RCTs)	The positive effects observed with external application of AV in preventing or treating infusion phlebitis compared with no intervention or external application of 33% or 50% $MgSO_4$. There is no strong evidence for preventing or treating infusion phlebitis with external application of AV.
Radiation-induced reactions							
Worthing-ton, 2011 (145)	preventing oral mucositis for patients with cancer	119 patients with head and neck cancer undergoing radiotherapy/ chemoradio-therapy	AV solution 8 weeks	placebo	mucositis event prevention	2 (RCTs)	● The meta-analysis showed a statistically significant benefit in favor of AV. ● Conclusions: there is weak unreliable evidence that AV may be beneficial in the prevention of moderate to severe mucositis.

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
Wound healing							
Norman, 2017(143)	burns	338 patients with burn wounds of any type, severity, extent or current infection status, in any age	topical AV (creams, gel or dressings with 1 study reported a 0.5% concentration) 14-60 days	topical antibiotics (SSD or framycetin cream)	● wound healing (mean time to wound healing, proportion of wounds healed) ● infection ● pain reduction	5 (RCTs)	● It is uncertain whether there is a difference in infection incidence, the mean time to healing, between AV and control group (no statistically difference reported from meta-analyses of 2 and 3 RCTs, respectively). ● 1 study reported unclear evidence in favor of AV in number of healing events compared with SSD. (RR 1.41, 95% CI 0.70 to 2.85). ● 1 study reported a slightly greater decrease in pain in the AV group compared with SSD group (MD 1.14, 95% CI 0.02 to 2.26).

Author, year	Disease/ indication	Population	Intervention (Type of AV) and duration	Comparison	Outcomes	N of included studies (design)	Main findings
Wang, 2013 (144)	acute surgical wound	97 patients with acute surgical wound (after skin biopsy, hemorrhoidectomy, surgical incision, and skin laser)	topical AV (cream, gel, fresh AV) 2-4 weeks	placebo, standard treatment, topical antibiotic	● wound healing (mean time to wound healing, proportion of wounds healed)	4 (RCTs)	● The meta-analysis of 2 RCTs showed higher proportion of patients with wounds healed in AV group with statistically significance; however, meta-analysis of another 2 RCTs showed no difference between the two groups in average wound healing time.
Wang, 2013 (144)	chronic wound	233 patients with pressure ulcer	topical AV (juice, gel, cream, and dressing) 21 days or until healing	standard treatment, topical antibiotic and topical disinfectant (0.5% iodine)	● wound healing (mean time to wound healing, proportion of wounds healed)	5 (RCTs)	● The meta-analysis of 5 RCTs showed higher proportion of wounds healed in AV group with statistically significance. ● 1 study reported no statistical different between AV and control group in the reduction of mean wound healing time.

Abbreviations: AV – *Aloe vera*, RCT – randomized controlled trial, q-RCT – quasi-randomized controlled trial, TA – triamcinolone acetonide, RR – relative risk or risk ratio, MD – mean difference, FPG – fasting plasma glucose, HbA1c – hemoglobin A1c, IV – intravenous, FBG – fasting blood glucose, HbA1C – Hemoglobin A1c, TC – total cholesterol, TG – triglyceride, LDL – low-density lipoprotein cholesterol, HDL – high- density lipoprotein cholesterol

Appendix 6 Detailed data of primary studies included in each meta-analysis that reported the continuous outcomes

Outcome examined	Primary studies		Comparator	Effect	Total	Control group			Intervention group			Effect	95%CI		SE
	Author	Year		metrics	n	n	mean	SD	n	mean	SD	size	LCI	UCI	
Dentistry															
Oral lichen planus (34)															
Pain and burning sensation	mansourian	2011	TA paste	WMD	46	23	0.75	0.08	23	0.81	0.08	0.06	0.01	0.11	0.02
Pain and burning sensation	reddy	2012	TA gel	WMD	20	10	2.15	2.6	10	1	1.83	-1.15	-3.12	0.82	1.01
Pain and burning sensation	salazar-sanchez	2010	placebo	WMD	55	24	3.7	3.3	31	2.5	3	-1.20	-2.89	0.49	0.86
Clinical improvement	mansourian	2011	TA paste	WMD	46	23	0.83	0.09	23	0.91	0.1	0.08	0.03	0.13	0.03
Clinical improvement	reddy	2012	TA gel	WMD	20	10	1.65	1.14	10	0.9	1.02	-0.75	-1.70	0.20	0.49
Clinical improvement	salazar-sanchez	2010	placebo	WMD	55	24	1.83	1.16	31	1.74	1.26	-0.09	-0.73	0.55	0.33
Oral submucous fibrosis (35)															
Burning sensation at 1 month	anuradha	2017	Intralesional steroids	WMD	74	37	5.27	1.04	37	3.52	1.12	-1.75	-2.24	-1.26	0.25
Burning sensation at 1 month	singh	2016	Antioxidant capsules	WMD	37	18	5.33	1.46	19	4.74	0.99	-0.59	-1.40	0.22	0.41
Burning sensation at 2 months	anuradha	2017	Intralesional steroids	WMD	74	37	3.42	1.85	37	1.8	0.65	-1.62	-2.25	-0.99	0.32
Burning sensation at 2 months	singh	2016	Antioxidant capsules	WMD	37	18	3.78	1.22	19	2.79	1.03	-0.99	-1.72	-0.26	0.37
Burning sensation at 3 months	anuradha	2017	Intralesional steroids	WMD	74	37	1.85	1.05	37	1.48	0.51	-0.37	-0.75	0.01	0.19

Outcome examined	Primary studies		Comparator	Effect	Total	Control group			Intervention group			Effect	95%CI		SE
	Author	Year		metrics	n	n	mean	SD	n	mean	SD	size	LCI	UCI	
Burning sensation at 3 months	singh	2016	Antioxidant capsules	WMD	37	18	2.06	1.16	19	0.53	0.49	-1.53	-2.11	-0.95	0.30
Burning sensation at 3 months	sudarshan	2012	Antioxidant capsules	WMD	20	10	19	17.3	10	15	23.2	-4.00	-21.94	13.94	9.15
Mouth opening at 1 month	anuradha	2017	Intralesional steroids	WMD	74	37	29.5	6.3	37	34.7	6.5	5.20	2.28	8.12	1.49
Mouth opening at 1 month	patil (a)	2015	lycopene	WMD	120	60	20.2	2.2	60	18.6	1.8	-1.60	-2.32	-0.88	0.37
Mouth opening at 1 month	patil (b)	2015	spirulina	WMD	42	21	20.9	2.8	21	20.4	2.2	-0.50	-2.02	1.02	0.78
Mouth opening at 1 month	patil	2014	oxitard capsule	WMD	120	60	21.6	2.6	60	18.6	1.8	-3.00	-3.80	-2.20	0.41
Mouth opening at 1 month	singh	2016	Antioxidant capsules	WMD	37	18	30.2	3.2	19	30.8	4.5	0.60	-1.91	3.11	1.28
Mouth opening at 2 months	anuradha	2017	Intralesional steroids	WMD	74	37	29.5	5.9	37	33.4	5.3	3.90	1.34	6.46	1.30
Mouth opening at 2 months	patil (a)	2015	lycopene	WMD	120	60	22.4	2.5	60	20.4	2	-2.00	-2.81	-1.19	0.41
Mouth opening at 2 months	patil (b)	2015	spirulina	WMD	42	21	23.4	2.2	21	22.1	1.5	-1.30	-2.44	-0.16	0.58
Mouth opening at 2 months	patil	2014	oxitard capsule	WMD	120	60	27.2	2.9	60	20.4	2	-6.80	-7.69	-5.91	0.46
Mouth opening at 2 months	singh	2016	Antioxidant capsules	WMD	37	18	30.9	3.3	19	31.8	4.5	0.90	-1.63	3.43	1.29

Outcome examined	Primary studies		Comparator	Effect	Total	Control group			Intervention group			Effect	95%CI		SE
	Author	Year		metrics	n	n	mean	SD	n	mean	SD	size	LCI	UCI	
Mouth opening at 3 months	anuradha	2017	Intralesional steroids	WMD	74	37	29.5	5.9	37	33.4	5.3	3.90	1.34	6.46	1.30
Mouth opening at 3 months	patil (a)	2015	lycopene	WMD	120	60	22.4	2.5	60	20.4	2	-2.00	-2.81	-1.19	0.41
mouth opening at 3 months	patil (b)	2015	spirulina	WMD	42	21	23.4	2.2	21	22.1	1.5	-1.30	-2.44	-0.16	0.58
Mouth opening at 3 months	patil	2014	oxitard capsule	WMD	120	60	27.2	2.9	60	20.4	2	-6.80	-7.69	-5.91	0.46
Mouth opening at 3 months	singh	2016	Antioxidant capsules	WMD	37	18	30.9	3.3	19	31.8	4.5	0.90	-1.63	3.43	1.29
Mouth opening at 3 months	sudarshan	2012	Antioxidant capsules	WMD	20	10	29.8	5.4	10	30.9	6.6	1.10	-4.19	6.39	2.70
Tongue protrusion at 1 month	anuradha	2017	Intralesional steroids	WMD	74	37	29.5	6.3	37	34.7	6.2	5.20	2.35	8.05	1.45
Tongue protrusion at 1 month	patil (a)	2015	lycopene	WMD	120	60	10.3	1.9	60	9.9	1.8	-0.40	-1.06	0.26	0.34
Tongue protrusion at 1 month	patil	2014	oxitard capsule	WMD	120	60	13.6	2.4	60	9.9	1.8	-3.70	-4.46	-2.94	0.39
Tongue protrusion at 1 month	singh	2016	Antioxidant capsules	WMD	37	18	43.9	3.5	19	43.1	7.2	-0.80	-4.42	2.82	1.85
Tongue protrusion at 2 months	anuradha	2017	Intralesional steroids	WMD	74	37	32.1	6.1	37	36.5	6.5	4.40	1.53	7.27	1.47
Tongue protrusion at 2 months	patil (a)	2015	lycopene	WMD	120	60	19.1	1.9	60	16.1	2.2	-3.00	-3.74	-2.26	0.38

Outcome examined	Primary studies		Comparator	Effect	Total	Control group			Intervention group			Effect	95%CI		SE
	Author	Year		metrics	n	n	mean	SD	n	mean	SD	size	LCI	UCI	
Tongue protrusion at 2 months	patil	2014	oxitard capsule	WMD	120	60	24.5	2.5	60	16.1	2.2	-8.40	-9.24	-7.56	0.43
Tongue protrusion at 2 months	singh	2016	Antioxidant capsules	WMD	37	18	44.7	3.6	19	44.6	7.6	-0.10	-3.90	3.70	1.94
Tongue protrusion at 3 months	anuradha	2017	Intralesional steroids	WMD	74	37	29.5	6.3	37	37.7	6.1	8.20	5.37	11.03	1.44
Tongue protrusion at 3 months	patil (a)	2015	lycopene	WMD	120	60	10.3	1.9	60	9.9	1.8	-0.40	-1.06	0.26	0.34
Tongue protrusion at 3 months	patil	2014	oxitard capsule	WMD	120	60	13.6	2.4	60	9.9	1.8	-3.70	-4.46	-2.94	0.39
Tongue protrusion at 3 months	singh	2016	Antioxidant capsules	WMD	37	18	43.9	3.5	19	43.1	7.2	-0.80	-4.42	2.82	1.85
Tongue protrusion at 3 months	sudarshan	2012	Antioxidant capsules	WMD	20	10	42.7	11.8	10	38.3	11.1	-4.40	-14.44	5.64	5.12
Cheek flexibility at 1 month	anuradha	2017	Intralesional steroids	WMD	74	37	0.41	0.09	37	0.3	0.06	-0.11	-0.14	-0.08	0.02
Cheek flexibility at 1 month	singh	2016	Antioxidant capsules	WMD	37	18	0.1	0.06	19	0.11	0.06	0.01	-0.03	0.05	0.02
Cheek flexibility at 2 months	anuradha	2017	Intralesional steroids	WMD	74	37	0.47	0.1	37	0.38	0.05	-0.09	-0.13	-0.05	0.02
Cheek flexibility at 2 months	singh	2016	Antioxidant capsules	WMD	37	18	0.14	0.11	19	0.16	0.09	0.02	-0.04	0.08	0.03
Cheek flexibility at 3 months	anuradha	2017	Intralesional steroids	WMD	74	37	0.53	0.1	37	0.43	0.05	-0.10	-0.14	-0.06	0.02

Outcome examined	Primary studies		Comparator	Effect	Total	Control group			Intervention group			Effect	95%CI		SE
	Author	Year		metrics	n	n	mean	SD	n	mean	SD	size	LCI	UCI	
Cheek flexibility at 3 months	singh	2016	Antioxidant capsules	WMD	37	18	0.18	0.16	19	0.19	0.17	0.01	-0.10	0.12	0.05
Cheek flexibility at 3 months	sudarshan	2012	Antioxidant capsules	WMD	20	10	0.14	0.05	10	0.18	0.06	0.04	-0.01	0.09	0.03
Wound healing															
Burns (38)															
Mean time to healing	Akhtar	1996	framycetin cream	WMD	100	50	30.9	0	50	18	0	exclude (not estimable)			
Mean time to healing	khorasani	2009	SSD cream	WMD	60	30	18.8	2.7	30	15.9	2	-2.90	-4.10	-1.70	0.61
Mean time to healing	Thamlikitkul	1991	SSD dressing	WMD	50	25	24.2	11.2	25	11	4.2	-13.20	-17.89	-8.51	2.39
Anti-diabetes															
Glucose lowering in prediabetic & early non-treated diabetic patients (42)															
FBG	Alinejad-Mofrad	2015	placebo	WMD	47	23	0	4.1	24	-7	4.2	-7.00	-9.37	-4.63	1.21
FBG	Choi	2013	placebo	WMD	122	62	-0.7	1.9	60	-3.1	1.5	-2.40	-3.01	-1.79	0.31
FBG	Choudhary	2014	placebo	WMD	60	30	-0.9	2.6	30	-15.8	2.7	-14.90	-16.24	-13.56	0.68
FBG	Devaraj	2013	placebo	WMD	29	15	3	14.1	14	-8	15.5	-11.00	-21.81	-0.19	5.52
FBG	Yongchaiyudha	1996	placebo	WMD	70	35	6.08	7.9	35	-108.4	6.6	-114.52	-117.93	-111.11	1.74
HbA1C	Alinejad-Mofrad	2015	placebo	WMD	47	23	0.02	0.2	24	-0.4	0.3	-0.42	-0.57	-0.27	0.07
HbA1C	Devaraj	2013	placebo	WMD	29	15	0	0.6	14	-0.3	0.6	-0.30	-0.74	0.14	0.22
Insulin level	Choi	2013	placebo	SMD	122	62	0.9	0.6	60	-1	0.7	-1.90	-2.13	-1.67	0.12

Outcome examined	Primary studies		Comparator	Effect	Total	Control group			Intervention group			Effect	95%CI		SE
	Author	Year		metrics	n	n	mean	SD	n	mean	SD	size	LCI	UCI	
Insulin level	Devaraj	2013	placebo	SMD	29	15	1	9.8	14	-4	9.8	-5.00	-12.14	2.14	3.64
Glucose lowering in type 2 diabetic patients (39)															
FBG	Liu	2002	no treatment	WMD	38	19	8.91	2.86	19	5.9	1.19	-3.01	-4.40	-1.62	0.71
FBG	Arora	2009	no treatment	WMD	24	12	8.51	0.52	12	6.9	1.84	-1.61	-2.69	-0.53	0.55
FBG	Huseini(a)	2012	placebo	WMD	60	30	10.62	2.38	30	9.32	1.82	-1.30	-2.37	-0.23	0.55
FBG	Huseini(b)	2012	placebo	WMD	70	35	10.57	2.67	35	9.4	2.05	-1.17	-2.29	-0.05	0.57
FBG	Zarrintan	2015	placebo	WMD	43	22	8.96	1.74	21	10.27	2.84	1.31	-0.11	2.73	0.72
HbA1C	Arora	2009	no treatment	WMD	24	12	66	13	12	44	11	-22.00	-31.64	-12.36	4.92
HbA1C	Huseini(a)	2012	placebo	WMD	50	30	62	20	20	49	12	-13.00	-21.88	-4.12	4.53
HbA1C	Huseini(b)	2012	placebo	WMD	53	35	61	18	18	52	11	-9.00	-16.83	-1.17	4.00
HbA1C	Zarrintan	2015	placebo	WMD	37	22	69	15	15	70	18	1.00	-10.06	12.06	5.64
Lipid lowering in prediabetic & early non-treated diabetic patients (42)															
TG	Alinejad-Mofrad	2015	placebo	WMD	47	23	0.6	1.4	24	-10.4	2.6	-11.00	-12.19	-9.81	0.61
TG	Choudhary	2014	placebo	WMD	60	30	-1.8	1.1	30	-13.6	1.7	-11.80	-12.52	-11.08	0.37
TG	Devaraj	2013	placebo	WMD	29	15	18.6	72.4	14	-8	83.7	-26.60	-83.74	30.54	29.15
TG	Yongchaiyudha	1996	placebo	WMD	70	35	18.13	10.2	35	-97.59	9.9	-115.72	-120.43	-111.01	2.40
TC	Alinejad-Mofrad	2015	placebo	WMD	47	23	1	3.3	24	-25	11.2	-26.00	-30.68	-21.32	2.39
TC	Choudhary	2014	placebo	WMD	60	30	-1.8	2.4	30	-13.4	2.5	-11.60	-12.84	-10.36	0.63

Outcome examined	Primary studies		Comparator	Effect	Total	Control group			Intervention group			Effect	95%CI		SE
	Author	Year		metrics	n	n	mean	SD	n	mean	SD	size	LCI	UCI	
TC	Devaraj	2013	placebo	WMD	29	15	4	29.6	14	-20	29	-24.00	-45.33	-2.67	10.89
TC	Yongchaiyudha	1996	placebo	WMD	70	35	13.41	6.7	35	0.69	6.5	-12.72	-15.81	-9.63	1.58
HDL	Alinejad-Mofrad	2015	placebo	WMD	47	23	-0.26	1.4	24	4.2	3.2	4.46	3.06	5.86	0.72
HDL	Choudhary	2014	placebo	WMD	60	30	0.6	0.9	30	2.3	0.9	1.70	1.24	2.16	0.23
HDL	Devaraj	2013	placebo	WMD	29	15	0.1	15.6	14	-2	10	-2.10	-11.57	7.37	4.83
LDL	Alinejad-Mofrad	2015	placebo	WMD	47	23	1	1.1	24	-14.1	1.3	-15.10	-15.79	-14.41	0.35
LDL	Choudhary	2014	placebo	WMD	60	30	-2	1.5	30	-13	1.9	-11.00	-11.87	-10.13	0.44
LDL	Devaraj	2013	placebo	WMD	29	15	1.1	21	14	-17	25.1	-18.10	-35.01	-1.19	8.63
Gastrointestinal disorders															
Irritate bowel syndrome (37)															
Symptom scores improvement	Davis	2006	placebo	SMD	49	23	13.74	85.03	26	39.12	77.45	25.38	-20.38	71.14	23.35
Symptom scores improvement	Hutchings	2011	placebo	SMD	25	13	2.49	1.7	12	3.5	2.26	1.01	-0.57	2.59	0.81
Symptom scores improvement	Storsrud	2015	placebo	SMD	63	31	23	73.1	32	58	76.35	35.00	-1.90	71.90	18.83

Outcome examined	Primary studies		Comparator	Effect	Total	Control group			Intervention group			Effect	95%CI		SE
	Author	Year		metrics	n	n	mean	SD	n	mean	SD	size	LCI	UCI	
Short-term IBS symptom score improvement (at 1 month)	Davis	2006	placebo	SMD	49	23	13.74	85.03	26	39.12	77.45	0.31	-0.25	0.88	0.29
Short-term IBS symptom score improvement (at 1 month)	Storsrud	2015	placebo	SMD	63	31	23	73.1	32	58	76.35	0.47	-0.03	0.97	0.26
Long-term IBS symptom score improvement (at 3 months)	Davis	2006	placebo	SMD	42	18	19.44	111.28	24	13.88	80.15	-0.06	-0.67	0.55	0.31
Long-term IBS symptom score improvement (at 3 months)	Hutchings	2011	placebo	SMD	25	13	2.8	1.91	12	4.08	2.64	0.56	-0.24	1.36	0.41

Abbreviations: n – number of participants, C – control group, I – intervention group (using *Aloe vera*), AV – *Aloe vera*, CI – confidence interval, SD – standard deviation, SE – standard error, LCI – Lower confidence interval, UCI – upper confidence interval, WMD – weight mean difference, SMD – standardized mean difference, CIP – chemotherapy induced phlebitis, OLP – oral lichen planus, OSF – oral submucous fibrosis, IBS – irritate bowel syndrome, FBG – fasting blood glucose, HbA1C – Hemoglobin A1c, TC – total cholesterol, TG – triglyceride, LDL – low-density lipoprotein cholesterol, HDL – high- density lipoprotein cholesterol

Appendix 7 Detailed data of primary studies included in each meta-analysis that reported the binary outcomes

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Gastrointestinal disorders															
Irritate bowel syndrome (11)															
Response rates	Davis	2006	placebo	RR	49	23	26	6	17	11	15	1.62	0.71	3.69	0.42
Response rates	Storsrud	2015	placebo	RR	63	31	32	11	20	18	14	1.59	0.90	2.79	0.29
Phlebitis															
Chemotherapy induced phlebitis (CIP) prevention (142)															
Overall incidence	Zhou	2001	conventional treatment	OR	66	36	30	7	29	0	30	0.06	1.18	1.48	0.00
Overall incidence	Dong	2008	conventional treatment	OR	57	28	29	9	19	4	25	0.34	1.26	0.67	0.09
Overall incidence	Xiao	2008	conventional treatment	OR	160	80	80	18	62	3	77	0.13	0.48	0.65	0.04
Overall incidence	Pan	2008	conventional treatment	OR	132	66	66	46	20	5	61	0.04	0.10	0.54	0.01
Overall incidence	Dong	2009	conventional treatment	OR	3000	1000	2000	285	715	65	1935	0.08	0.11	0.14	0.06
Overall incidence	Chen	2010	conventional treatment	OR	78	38	40	18	20	6	34	0.20	0.58	0.55	0.07
Overall incidence	Yang	2010	conventional treatment	OR	186	90	96	20	70	5	91	0.19	0.54	0.53	0.07

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Overall incidence	Wan	2011	conventional treatment	OR	106	53	53	33	20	8	45	0.11	0.27	0.48	0.04
Overall incidence	Lei	2012	conventional treatment	OR	98	49	49	30	19	8	41	0.12	0.32	0.49	0.05
Overall incidence	Chen	2012	conventional treatment	OR	100	50	50	10	40	5	45	0.44	1.41	0.59	0.14
Incidence of 1st-degree CIP	Xiao	2008	conventional treatment	OR	160	80	80	7	73	3	77	0.41	1.63	0.71	0.10
Incidence of 1st-degree CIP	Dong	2009	conventional treatment	OR	3000	1000	2000	12	988	42	1958	1.77	3.37	0.33	0.93
Incidence of 1st-degree CIP	Lei	2012	conventional treatment	OR	98	49	49	10	39	4	45	0.35	1.19	0.63	0.10
Incidence of 1st-degree CIP	Chen	2010	conventional treatment	OR	186	90	96	4	86	5	91	1.18	4.55	0.69	0.31
Incidence of 1st-degree CIP	Pan	2008	conventional treatment	OR	132	66	66	15	51	0	66	0.02	0.43	1.45	0.00
Incidence of 1st-degree CIP	Wan	2011	conventional treatment	OR	106	53	53	0	53	6	47	14.64	266.87	1.48	0.80
Incidence of 1st-degree CIP	Dong	2008	conventional treatment	OR	57	28	29	9	19	3	26	0.24	1.02	0.73	0.06
Incidence of 1st-degree CIP	Yang	2010	conventional treatment	OR	186	90	96	18	72	5	91	0.22	0.62	0.53	0.08

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Incidence of 2nd-degree CIP	Xiao	2008	conventional treatment	OR	160	80	80	4	76	1	79	0.24	2.20	1.13	0.03
Incidence of 2nd-degree CIP	Dong	2009	conventional treatment	OR	3000	1000	2000	125	875	23	1977	0.08	0.13	0.23	0.05
Incidence of 2nd-degree CIP	Lei	2012	conventional treatment	OR	98	49	49	13	36	3	46	0.18	0.68	0.68	0.05
Incidence of 2nd-degree CIP	Chen	2010	conventional treatment	OR	186	90	96	5	85	1	95	0.18	1.56	1.11	0.02
Incidence of 2nd-degree CIP	Pan	2008	conventional treatment	OR	132	66	66	31	35	4	62	0.07	0.22	0.57	0.02
Incidence of 2nd-degree CIP	Wan	2011	conventional treatment	OR	106	53	53	7	46	1	52	0.13	1.07	1.09	0.01
Incidence of 2nd-degree CIP	Dong	2008	conventional treatment	OR	57	28	29	3	25	1	28	0.30	3.05	1.19	0.03
Incidence of 2nd-degree CIP	Yang	2010	conventional treatment	OR	186	90	96	4	86	0	96	0.10	1.88	1.50	0.01
Incidence of 3rd-degree CIP	Xiao	2008	conventional treatment	OR	160	80	80	2	78	0	80	0.20	4.13	1.56	0.01
Incidence of 3rd-degree CIP	Dong	2009	conventional treatment	OR	3000	1000	2000	68	932	0	2000	0.00	0.06	1.42	0.00
Incidence of 3rd-degree CIP	Lei	2012	conventional treatment	OR	98	49	49	5	44	1	48	0.18	1.63	1.12	0.02

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Incidence of 3rd-degree CIP	Chen	2010	conventional treatment	OR	186	90	96	3	87	0	96	0.13	2.54	1.52	0.01
Incidence of 3rd-degree CIP	Pan	2008	conventional treatment	OR	132	66	66	8	58	1	65	0.11	0.92	1.08	0.01
Incidence of 3rd-degree CIP	Wan	2011	conventional treatment	OR	106	53	53	8	45	1	52	0.11	0.90	1.08	0.01
Incidence of 3rd-degree CIP	Dong	2008	conventional treatment	OR	57	28	29	1	27	0	29	0.31	7.95	1.65	0.01
Incidence of 3rd-degree CIP	Yang	2010	conventional treatment	OR	186	90	96	1	89	0	96	0.31	7.69	1.64	0.01
Chemotherapy induced phlebitis-treatment (142)															
Overall efficacy rate	Dong	2001	50% $MgSO_4$	RR	154	74	80	53	21	75	5	1.31	1.53	0.08	1.12
Overall efficacy rate	Gao	2006	50% $MgSO_4$	RR	100	50	50	37	13	48	2	1.30	1.54	0.09	1.09
Overall efficacy rate	Yang	2008	50% $MgSO_4$	RR	100	48	52	39	9	50	2	1.18	1.37	0.08	1.02
Overall efficacy rate	Deng	2010	50% $MgSO_4$	RR	85	42	43	30	12	43	0	1.39	1.69	0.10	1.15
Overall efficacy rate	Tang	2011	50% $MgSO_4$	RR	66	30	36	22	8	34	2	1.29	1.62	0.12	1.02
Overall efficacy rate	Zhang	2015	50% $MgSO_4$	RR	42	21	21	16	5	20	1	1.25	1.62	0.13	0.97
Overall cure rate	Deng	2010	50% $MgSO_4$	RR	85	42	43	4	38	36	7	8.79	22.54	0.48	3.43
Overall cure rate	Tang	2011	50% $MgSO_4$	RR	66	30	36	13	17	26	10	1.67	2.63	0.23	1.06
Overall cure rate	Zhang	2015	50% $MgSO_4$	RR	42	21	21	6	15	10	11	1.67	3.75	0.41	0.74
Prevention of Infusion phlebitis (14)															

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Total incidence of phlebitis (duration used 5 days)	Peng	2009	no treatment	RR	50	25	25	15	10	18	7	1.20	1.80	0.21	0.80
Total incidence of phlebitis (duration used 5 days)	Xiao	2008	no treatment	RR	160	80	80	18	62	3	77	0.17	0.54	0.60	0.05
Total incidence of phlebitis (duration used 5 days)	Yao	2009	no treatment	RR	196	98	98	98	0	4	94	0.05	0.11	0.46	0.02
Total incidence of phlebitis (duration used 5 days)	Zheng	2010	no treatment	RR	50	25	25	18	7	3	22	0.17	0.50	0.56	0.06
Total incidence of phlebitis (duration used 1-7 days)	Wang	2006	no treatment	RR	220	110	110	54	56	20	90	0.37	0.57	0.22	0.24
Total incidence of phlebitis (duration used 1-7 days)	Yu	2006	no treatment	RR	150	75	75	33	42	7	68	0.21	0.45	0.38	0.10
Total incidence of phlebitis (duration used 3 days)	Ji	2007	no treatment	RR	140	70	70	29	41	6	64	0.21	0.47	0.42	0.09

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Total incidence of phlebitis (duration used 3 days)	Liu	2006	no treatment	RR	86	43	43	35	8	18	25	0.51	0.75	0.19	0.35
Total incidence of phlebitis (duration used 3 days)	Liu(b)	2012	no treatment	RR	86	43	43	35	8	18	25	0.51	0.75	0.19	0.35
Total incidence of phlebitis (duration used 2-3 days)	Dong	2008	no treatment	RR	57	28	29	9	19	4	25	0.43	1.24	0.54	0.15
Total incidence of phlebitis (duration used 2-3 days)	Pan	2008	no treatment	RR	132	66	66	46	20	5	61	0.11	0.26	0.44	0.05
Incidence of 2nd-degree phlebitis	Li	2011	no treatment	RR	80	40	40	9	31	2	38	0.22	0.96	0.75	0.05
Incidence of 2nd-degree phlebitis	Cao	2008	no treatment	RR	76	38	38	1	37	0	38	0.33	7.93	1.62	0.01
Incidence of 2nd-degree phlebitis	Xiao	2008	no treatment	RR	160	80	80	6	74	1	79	0.17	1.35	1.07	0.02
Incidence of 2nd-degree phlebitis	Yao	2009	no treatment	RR	196	98	98	9	89	2	96	0.22	1.00	0.77	0.05
Incidence of 2nd-degree phlebitis	Zheng	2010	no treatment	RR	50	25	25	5	20	0	25	0.09	1.56	1.45	0.01

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Incidence of 2nd-degree phlebitis	Wang	2006	no treatment	RR	220	110	110	32	78	8	102	0.25	0.52	0.37	0.12
Incidence of 2nd-degree phlebitis	Yu	2006	no treatment	RR	150	75	75	17	58	2	73	0.12	0.49	0.73	0.03
Incidence of 2nd-degree phlebitis	Ji	2007	no treatment	RR	142	72	70	14	58	2	68	0.15	0.62	0.74	0.03
Incidence of 2nd-degree phlebitis	Liu	2006	no treatment	RR	86	43	43	17	26	8	35	0.47	0.97	0.37	0.23
Incidence of 2nd-degree phlebitis	Liu(b)	2012	no treatment	RR	86	43	43	10	33	6	37	0.60	1.51	0.47	0.24
Incidence of 2nd-degree phlebitis	Dong	2008	no treatment	RR	57	28	29	5	23	1	28	0.19	1.55	1.06	0.02
Incidence of 2nd-degree phlebitis	Pan	2008	no treatment	RR	132	66	66	39	27	1	65	0.03	0.18	1.00	0.00
Incidence of 2nd-degree phlebitis	Tan	2002	no treatment	RR	150	76	74	56	20	10	64	0.18	0.33	0.30	0.10
Incidence of 2nd-degree phlebitis	Hu	2009	no treatment	RR	3000	1000	2000	193	807	23	1977	0.06	0.09	0.22	0.04
Incidence of 2nd-degree phlebitis (duration used 5 days)	Cao	2008	no treatment	RR	76	38	38	1	37	0	38	0.33	7.93	1.62	0.01

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Incidence of 2nd-degree phlebitis (duration used 5 days)	Xiao	2008	no treatment	RR	160	80	80	6	74	1	79	0.17	1.35	1.07	0.02
Incidence of 2nd-degree phlebitis (duration used 5 days)	Yao	2009	no treatment	RR	196	98	98	9	89	2	96	0.22	1.00	0.77	0.05
Incidence of 2nd-degree phlebitis (duration used 5 days)	Zheng	2010	no treatment	RR	50	25	25	5	20	0	25	0.09	1.56	1.45	0.01
Incidence of 2nd-degree phlebitis (duration used 1-7 days)	Wang	2006	no treatment	RR	220	110	110	32	78	8	102	0.25	0.52	0.37	0.12
Incidence of 2nd-degree phlebitis (duration used 1-7 days)	Yu	2006	no treatment	RR	150	75	75	17	58	2	73	0.12	0.49	0.73	0.03
Incidence of 2nd-degree phlebitis	Ji	2007	no treatment	RR	142	72	70	14	58	2	68	0.15	0.62	0.74	0.03

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
(duration used 3 days)															
Incidence of 2nd-degree phlebitis (duration used 3 days)	Liu	2006	no treatment	RR	86	43	43	17	26	8	35	0.47	0.97	0.37	0.23
Incidence of 2nd-degree phlebitis (duration used 3 days)	Liu(b)	2012	no treatment	RR	86	43	43	10	33	6	37	0.60	1.51	0.47	0.24
Incidence of 2nd-degree phlebitis (duration used 2-3 days)	Dong	2008	no treatment	RR	57	28	29	5	23	1	28	0.19	1.55	1.06	0.02
Incidence of 2nd-degree phlebitis (duration used 2-3 days)	Pan	2008	no treatment	RR	132	66	66	39	27	1	65	0.03	0.18	1.00	0.00
Incidence of 3rd-degree phlebitis	Li	2011	no treatment	RR	80	40	40	2	38	1	39	0.50	5.30	1.20	0.05
Incidence of 3rd-degree phlebitis	Cao	2008	no treatment	RR	76	38	38	0	38	0	38	NA	NA	NA	NA

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Incidence of 3rd-degree phlebitis	Xiao	2008	no treatment	RR	160	80	80	2	78	0	80	0.20	4.10	1.54	0.01
Incidence of 3rd-degree phlebitis	Yao	2009	no treatment	RR	196	98	98	0	98	0	98	NA	NA	NA	NA
Incidence of 3rd-degree phlebitis	Zheng	2010	no treatment	RR	50	25	25	1	24	0	25	0.33	7.81	1.61	0.01
Incidence of 3rd-degree phlebitis	Wang	2006	no treatment	RR	220	110	110	8	102	1	109	0.13	0.98	1.05	0.02
Incidence of 3rd-degree phlebitis	Yu	2006	no treatment	RR	150	75	75	6	69	0	75	0.08	1.34	1.46	0.00
Incidence of 3rd-degree phlebitis	Ji	2007	no treatment	RR	142	72	70	1	71	0	70	0.34	8.27	1.62	0.01
Incidence of 3rd-degree phlebitis	Liu	2006	no treatment	RR	86	43	43	7	36	2	41	0.29	1.30	0.77	0.06
Incidence of 3rd-degree phlebitis	Liu(b)	2012	no treatment	RR	86	43	43	7	36	2	41	0.29	1.30	0.77	0.06
Incidence of 3rd-degree phlebitis	Dong	2008	no treatment	RR	57	28	29	2	26	0	29	0.19	3.86	1.53	0.01
Incidence of 3rd-degree phlebitis	Pan	2008	no treatment	RR	132	66	66	8	58	0	66	0.06	1.00	1.45	0.00
Incidence of 3rd-degree phlebitis	Tan	2002	no treatment	RR	150	76	74	24	52	0	74	0.02	0.34	1.42	0.00

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Incidence of 3rd-degree phlebitis	Hu	2009	no treatment	RR	3000	1000	2000	68	932	0	2000	0.00	0.06	1.42	0.00
Incidence of 3rd-degree phlebitis (duration used 5 days)	Cao	2008	no treatment	RR	76	38	38	0	38	0	38	NA	NA	NA	NA
Incidence of 3rd-degree phlebitis (duration used 5 days)	Xiao	2008	no treatment	RR	160	80	80	2	78	0	80	0.20	4.10	1.54	0.01
Incidence of 3rd-degree phlebitis (duration used 5 days)	Yao	2009	no treatment	RR	196	98	98	0	98	0	98	NA	NA	NA	NA
Incidence of 3rd-degree phlebitis (duration used 5 days)	Zheng	2010	no treatment	RR	50	25	25	1	24	0	25	0.33	7.81	1.61	0.01
Incidence of 3rd-degree phlebitis (duration used 1-7 days)	Wang	2006	no treatment	RR	220	110	110	8	102	1	109	0.13	0.98	1.05	0.02

Outcome examined	Primary studies		Comparator	Effect metrics	Total n	n/group		Control (C)		Intervention (I)		Effect size	95%CI		SE
	Author	Year				C	I	event	no event	event	no event		LCI	UCI	
Incidence of 3rd-degree phlebitis (duration used 1-7 days)	Yu	2006	no treatment	RR	150	75	75	6	69	0	75	0.08	1.34	1.46	0.00
Incidence of 3rd-degree phlebitis (duration used 3 days)	Ji	2007	no treatment	RR	142	72	70	1	71	0	70	0.34	8.27	1.62	0.01
Incidence of 3rd-degree phlebitis (duration used 3 days)	Liu	2006	no treatment	RR	86	43	43	7	36	2	41	0.29	1.30	0.77	0.06
Incidence of 3rd-degree phlebitis (duration used 3 days)	Liu(b)	2012	no treatment	RR	86	43	43	7	36	2	41	0.29	1.30	0.77	0.06
Total incidence of phlebitis	Wang	2006	potato slice	RR	210	100	110	19	81	20	90	0.96	1.69	0.29	0.54
Total incidence of phlebitis	Wu	2009	potato slice	RR	66	33	33	12	21	10	23	0.83	1.66	0.35	0.42
Total incidence of 2nd-degree phlebitis	Wang	2006	potato slice	RR	210	100	110	6	94	8	102	1.21	3.37	0.52	0.44

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Total incidence of 2nd-degree phlebitis	Wu	2009	potato slice	RR	66	33	33	3	30	3	30	1.00	4.60	0.78	0.22
Total incidence of phlebitis	Chen	2012	33% MgSO4	RR	100	50	50	10	40	5	45	0.50	1.36	0.51	0.18
Total incidence of phlebitis	Hou	2010	33% MgSO4	RR	100	50	50	20	30	8	42	0.40	0.82	0.37	0.19
Total incidence of phlebitis	Ren	2008	50% $MgSO_4$	RR	153	49	104	12	37	6	98	0.24	0.59	0.47	0.09
Total incidence of phlebitis	Zhang	2010	50% $MgSO_4$	RR	95	63	32	63	0	19	13	0.60	0.79	0.15	0.45
Total incidence of phlebitis	Ren	2008	50% $MgSO_4$	RR	155	50	105	12.5	37.5	6.5	98.5	0.24	0.59	0.47	0.09
Total incidence of phlebitis	Zhang	2010	50% $MgSO_4$	RR	97	64	33	63.5	0.5	19.5	13.5	0.60	0.79	0.15	0.45
Total incidence of 2nd-degree phlebitis	Ren	2008	50% $MgSO_4$	RR	153	49	104	6	43	3	101	0.24	0.90	0.69	0.06
Total incidence of 2nd-degree phlebitis	Zhang	2010	50% $MgSO_4$	RR	95	63	32	40	23	6	26	0.30	0.62	0.38	0.14
Total incidence of 2nd-degree phlebitis	Ren	2008	50% $MgSO_4$	RR	153	49	104	4	45	2	102	0.24	1.24	0.85	0.04
Total incidence of 2nd-degree phlebitis	Zhang	2010	50% $MgSO_4$	RR	95	63	32	6	57	1	31	0.33	2.61	1.06	0.04

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Treatment of Infusion phlebitis (14)															
Rate of resolution of phlebitis: marked improvement	Dong	2001	33% $MgSO_4$	RR	154	74	80	16	58	26	54	1.50	2.57	0.27	0.88
Rate of resolution of phlebitis: marked improvement	Li	2006	33% $MgSO_4$	RR	148	68	80	21	47	56	24	2.27	3.33	0.20	1.54
Rate of resolution of phlebitis: total improvement	Dong	2001	33% $MgSO_4$	RR	154	74	80	53	21	75	5	1.31	1.53	0.08	1.12
Rate of resolution of phlebitis: total improvement	Gao	2012	33% $MgSO_4$	RR	120	60	60	51	9	58	2	1.14	1.28	0.06	1.01
Rate of resolution of phlebitis: total improvement	Li	2006	33% $MgSO_4$	RR	148	68	80	60	8	80	0	1.13	1.24	0.05	1.04
Rate of resolution of phlebitis: recovery	Deng	2012	50% $MgSO_4$	RR	116	56	60	30	26	47	13	1.46	1.93	0.14	1.11
Rate of resolution of phlebitis: recovery	Gao	2007	50% $MgSO_4$	RR	90	42	48	30	12	39	9	1.14	1.44	0.12	0.90

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Rate of resolution of phlebitis: recovery	Li	2003	50% $MgSO_4$	RR	88	42	46	13	29	28	18	1.97	3.27	0.26	1.18
Rate of resolution of phlebitis: recovery	Yang	2008	50% $MgSO_4$	RR	100	48	52	15	33	30	22	1.85	2.98	0.25	1.14
Rate of resolution of phlebitis: recovery	Liu	2012	50% $MgSO_4$	RR	50	25	25	4	21	14	11	3.50	9.17	0.49	1.34
Rate of resolution of phlebitis: recovery	Deng	2010	50% $MgSO_4$	RR	85	42	43	26	16	36	7	1.35	1.77	0.14	1.03
Rate of resolution of phlebitis: recovery	Tang	2011	50% $MgSO_4$	RR	66	30	36	13	17	26	10	1.67	2.63	0.23	1.06
Rate of resolution of phlebitis: recovery	Deng	2012	50% $MgSO_4$	RR	116	56	60	30	26	47	13	1.46	1.93	0.14	1.11
Rate of resolution of phlebitis: recovery	Gao	2007	50% $MgSO_4$	RR	90	42	48	30	12	39	9	1.14	1.44	0.12	0.90
Rate of resolution of phlebitis: recovery	Li	2003	50% $MgSO_4$	RR	88	42	46	13	29	28	18	1.97	3.27	0.26	1.18
Rate of resolution of phlebitis: recovery	Yang	2008	50% $MgSO_4$	RR	100	48	52	15	33	30	22	1.85	2.98	0.25	1.14
Rate of resolution of phlebitis: recovery	Deng	2010	50% $MgSO_4$	RR	85	42	43	26	16	36	7	1.35	1.77	0.14	1.03
Rate of resolution of phlebitis: recovery	Tang	2011	50% $MgSO_4$	RR	66	30	36	13	17	26	10	1.67	2.63	0.23	1.06

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Rate of resolution of phlebitis: marked improvement	Chen	2010	50% $MgSO_4$	RR	100	50	50	3	47	38	12	12.67	38.36	0.57	4.18
Rate of resolution of phlebitis: marked improvement	Deng	2012	50% $MgSO_4$	RR	116	56	60	42	14	58	2	1.29	1.51	0.08	1.10
Rate of resolution of phlebitis: marked improvement	Gao	2007	50% $MgSO_4$	RR	90	42	48	34	8	45	3	1.16	1.36	0.08	0.98
Rate of resolution of phlebitis: marked improvement	Li	2003	50% $MgSO_4$	RR	88	42	46	26	16	41	5	1.44	1.86	0.13	1.11
Rate of resolution of phlebitis: marked improvement	Li	2009	50% $MgSO_4$	RR	120	60	60	26	34	36	24	1.38	1.98	0.18	0.97
Rate of resolution of phlebitis: marked improvement	Wang	2010	50% $MgSO_4$	RR	65	32	33	9	23	26	7	2.80	5.01	0.30	1.57
Rate of resolution of phlebitis: marked improvement	Yang	2008	50% $MgSO_4$	RR	100	48	52	26	22	45	7	1.60	2.12	0.14	1.21

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Rate of resolution of phlebitis: marked improvement	Liu	2012	50% $MgSO_4$	RR	50	25	25	9	16	20	5	2.22	3.88	0.29	1.27
Rate of resolution of phlebitis: marked improvement	Deng	2010	50% $MgSO_4$	RR	85	42	43	30	12	43	0	1.39	1.69	0.10	1.15
Rate of resolution of phlebitis: marked improvement	Chen	2010	50% $MgSO_4$	RR	100	50	50	3	47	38	12	12.67	38.36	0.57	4.18
Rate of resolution of phlebitis: marked improvement	Deng	2012	50% $MgSO_4$	RR	116	56	60	42	14	58	2	1.29	1.51	0.08	1.10
Rate of resolution of phlebitis: marked improvement	Gao	2007	50% $MgSO_4$	RR	90	42	48	34	8	45	3	1.16	1.36	0.08	0.98
Rate of resolution of phlebitis: marked improvement	Li	2003	50% $MgSO_4$	RR	88	42	46	26	16	41	5	1.44	1.86	0.13	1.11
Rate of resolution of phlebitis: marked improvement	Li	2009	50% $MgSO_4$	RR	120	60	60	26	34	36	24	1.38	1.98	0.18	0.97
Rate of resolution of	Wang	2010	50% $MgSO_4$	RR	65	32	33	9	23	26	7	2.80	5.01	0.30	1.57

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
phlebitis: marked improvement															
Rate of resolution of phlebitis: marked improvement	Yang	2008	50% $MgSO_4$	RR	100	48	52	26	22	45	7	1.60	2.12	0.14	1.21
Rate of resolution of phlebitis: total improvement	Chen	2010	50% $MgSO_4$	RR	100	50	50	36	14	50	0	1.38	1.65	0.09	1.16
Rate of resolution of phlebitis: total improvement	Deng	2012	50% $MgSO_4$	RR	116	56	60	49	7	60	0	1.14	1.27	0.05	1.03
Rate of resolution of phlebitis: total improvement	Gao	2007	50% $MgSO_4$	RR	90	42	48	36	6	47	1	1.14	1.30	0.07	1.00
Rate of resolution of phlebitis: total improvement	Li	2003	50% $MgSO_4$	RR	88	42	46	34	8	45	1	1.21	1.41	0.08	1.04
Rate of resolution of phlebitis: total improvement	Li	2009	50% $MgSO_4$	RR	120	60	60	44	16	54	6	1.23	1.46	0.09	1.03
Rate of resolution of phlebitis:	Wang	2010	50% $MgSO_4$	RR	65	32	33	25	7	33	0	1.28	1.54	0.10	1.05

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
total improvement															
Rate of resolution of phlebitis: total improvement	Yang	2008	50% $MgSO_4$	RR	100	48	52	39	9	50	2	1.18	1.37	0.08	1.02
Rate of resolution of phlebitis: total improvement	Liu	2012	50% $MgSO_4$	RR	50	25	25	15	10	23	2	1.53	2.15	0.17	1.09
Rate of resolution of phlebitis: total improvement	Deng	2010	50% $MgSO_4$	RR	85	42	43	30	12	43	0	1.39	1.69	0.10	1.15
Rate of resolution of phlebitis: total improvement	Tang	2011	50% $MgSO_4$	RR	66	30	36	22	8	34	2	1.29	1.62	0.12	1.02
Rate of resolution of phlebitis: total improvement at 3 days	Chen	2010	50% $MgSO_4$	RR	100	50	50	36	14	50	0	1.38	1.65	0.09	1.16
Rate of resolution of phlebitis: total improvement at 3 days	Deng	2012	50% $MgSO_4$	RR	116	56	60	49	7	60	0	1.14	1.27	0.05	1.03
Rate of resolution of	Gao	2007	50% $MgSO_4$	RR	90	42	48	36	6	47	1	1.14	1.30	0.07	1.00

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
phlebitis: total improvement at 3 days															
Rate of resolution of phlebitis: total improvement at 3 days	Li	2003	50% $MgSO_4$	RR	88	42	46	34	8	45	1	1.21	1.41	0.08	1.04
Rate of resolution of phlebitis: total improvement at 3 days	Li	2009	50% $MgSO_4$	RR	120	60	60	44	16	54	6	1.23	1.46	0.09	1.03
Rate of resolution of phlebitis: total improvement at 3 days	Wang	2010	50% $MgSO_4$	RR	65	32	33	25	7	33	0	1.28	1.54	0.10	1.05
Rate of resolution of phlebitis: total improvement at 3 days	Yang	2008	50% $MgSO_4$	RR	100	48	52	39	9	50	2	1.18	1.37	0.08	1.02
Rate of resolution of phlebitis: total improvement	Deng	2010	50% $MgSO_4$	RR	85	42	43	30	12	43	0	1.39	1.69	0.10	1.15

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
Rate of resolution of phlebitis: total improvement	Tang	2011	50% $MgSO_4$	RR	66	30	36	22	8	34	2	1.29	1.62	0.12	1.02
Rate of resolution of phlebitis: recovery	Jin	2010	Hirudoid	RR	63	31	32	10	21	25	7	2.42	4.16	0.28	1.41
Rate of resolution of phlebitis: recovery	Zhang	2013	Hirudoid	RR	160	80	80	27	53	52	28	1.93	2.72	0.18	1.36
Rate of resolution of phlebitis: recovery	Zhong	2011	Sulphanilamide	RR	60	30	30	19	11	25	5	1.32	1.80	0.16	0.96
Rate of resolution of phlebitis: marked improvement	Chen	2009	Sulphonic acid mucopolysaccharide	RR	40	20	20	12	8	15	5	1.25	1.94	0.22	0.81
Rate of resolution of phlebitis: marked improvement	Jin	2010	Hirudoid	RR	63	31	32	19	12	28	4	1.43	1.94	0.16	1.05
Rate of resolution of phlebitis: marked improvement	Zhong	2011	Sulphanilamide	RR	60	30	30	24	6	29	1	1.21	1.46	0.10	1.00
Rate of resolution of	Jin	2010	Hirudoid	RR	63	31	32	22	9	31	1	1.37	1.72	0.12	1.08

Outcome examined	Primary studies		Comparator	Effect	Total	n/group		Control (C)		Intervention (I)		Effect	95%CI		SE
	Author	Year		metrics	n	C	I	event	no event	event	no event	size	LCI	UCI	
phlebitis: total improvement															
Rate of resolution of phlebitis: total improvement	Zhong	2011	Sulphanilamide	RR	60	30	30	24	6	29	1	1.21	1.46	0.10	1.00
Rate of resolution of phlebitis: total improvement	Chen	2009	Sulphonic acid mucopolysaccharide	RR	40	20	20	18	2	19	1	1.06	1.26	0.09	0.88
Rate of resolution of phlebitis: total improvement	Zhang	2013	Hirudoid	RR	160	80	80	58	22	77	3	1.33	1.53	0.07	1.15
Radiation-induced reaction															
Radiation-induced mucositis (145)															
Mucositis prevention	Su	2004	placebo	RR	58	30	28	21	9	18	10	0.92	1.32	0.19	0.64
Mucositis prevention	Puata weepong	2009	placebo	RR	61	31	30	27	4	16	14	0.61	0.88	0.18	0.43
Wound healing															
Burns (143)															
Infection	khorasani	2009	SSD	RR	60	30	30	0	30	0	30	NA	NA	NA	NA
Infection	Panahi	2012	SSD 1% cream	RR	111	55	56	0	55	1	56	2.95	70.82	1.62	0.12
Infection	Shahzad	2013	SSD 1% cream	RR	50	25	25	4	21	3	22	0.75	3.01	0.71	0.19

Abbreviation: n – number of participants, C – control group, I – intervention group (using *Aloe vera*), I^2 – heterogeneity, CI – confidence interval, LCI – Lower confidence interval, UCI- upper confidence interval, PI – prediction interval, LPI – lower prediction interval, UPI – upper prediction interval, SE – standard error, ES – effect size, WMD –weight mean difference, SMD – standardized mean difference, OR – odds ratio, RR – relative risk, NA– not applicable, NR-not reported, SSD-silver sulfadiazine, $MgSO_4$ – magnesium sulfate

Appendix 8 Statistical analysis: command used in the STATA program

8.1. Effect size and heterogeneity (I^2)

Type of outcome variables	Command used in STATA program
Continuous outcomes **(i.e. mean difference)**	Mean difference: metan tsample tmean tsd csample cmean csd, nostandard random rfdist rflevel(95) textsize(200) label (namevar = author, yearvar = year)
	Standardized mean difference: metan tsample tmean tsd csample cmean csd, random rfdist rflevel(95) textsize(200) label (namevar = author, yearvar = year)
Binary outcomes **(i.e. relative risk, odd ratio)**	Relative risk: metan tevent tnonevent cevent cnonevent, random rfdist rflevel(95) textsize(200) label (namevar = author, yearvar = year)
	Odd ratio: metan tevent tnonevent cevent cnonevent, or random rfdist rflevel(95) textsize(200) label (namevar = author, yearvar = year)

8.2 Prediction interval

8.2.1 Using similar commands as for the effect size calculation

8.2.2 Prediction interval can be obtained from the forest plot graph. As shown in the box in figure 5. However, if the input data came from ≤ 2 studies, prediction can't be calculated.

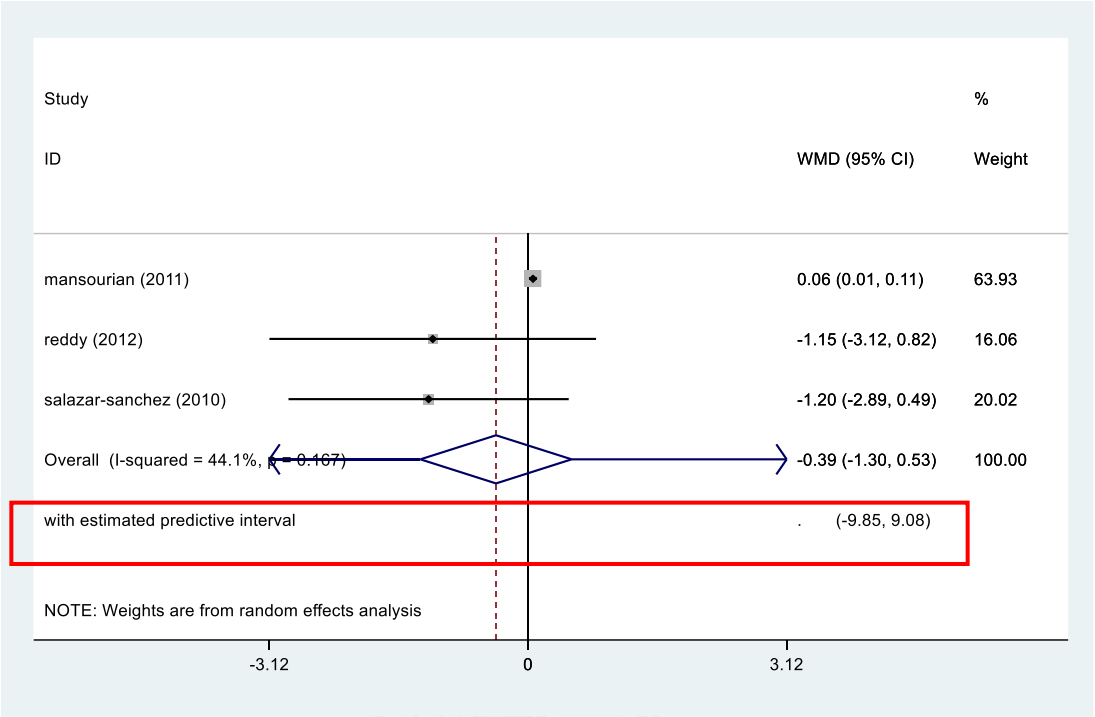

Figure 5 Example of prediction interval in the forest plot

8.3 Small study effects

Small study effects in this umbrella review were estimated using Egger's test. The main command used in STATA for egger's test was *metabias varlist [if] [in], egger.* The default type of variable assigned in STATA program is OR with a confident interval at level 95. As in the *metan* command, *varlist* should contain either four or two variables as follows;

- When four variables are given, these are assumed to be cell counts for the 2 x 2 table in this order: cases and noncases for the experimental group, then cases and noncases for the control group (d1 h1d0 h0). Then, the command 'metabias tevent tnonevent cevent cnonevent, egger' or 'metabias _ES _seES, egger' can be used.
- When two variables are specified, these are assumed to be the effect estimate and its standard error (theta se_theta). It is recommended that ratio-based effect estimates are log transformed as in *metan*. Then, the

command ‘metabias theta se_theta, egger’ or ‘metabias _ES _seES, egger’ can be used.

8.4 Excess significance test

The excess statistical significance test was performed to evaluate whether the number of positive studies among those in a meta-analysis is too large based on the power that these studies have to detect plausible effects at α equal to 0.05. Steps of performing the excess significance test using STATA program are described below;

4.1. power calculation:

4.1.1 Select the power and sample-size analysis from the statistics button on the menu bar.

4.1.2 Select an organizing method from type of outcome or analysis type. For example, when the meta-analysis compared two independent proportions, the method should be selected as shown in figure 6.

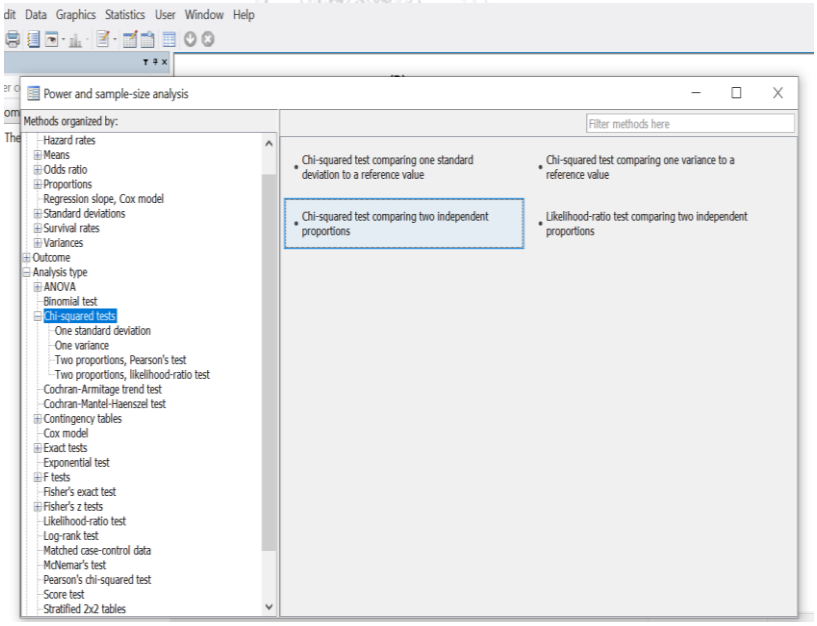

Figure 6 Selection of the Power calculation from menu bar in STATA program

4.1.3 Fill data into all the blanks (e.g. total sample size, treatment sample size with allocation ratio, control sample size, and number of cases reported in control for each study reported in a meta-analysis.), as shown in figure 3. The

plausible effect size for power calculation, to be filled in the highlighted box, can be obtained from the largest study's effect size of each meta-analysis. For example, if the largest study in the meta-analysis reported effect size (OR) of 0.8 (0.7-0.9), 0.8 will be considered as a plausible effect size for power calculation.

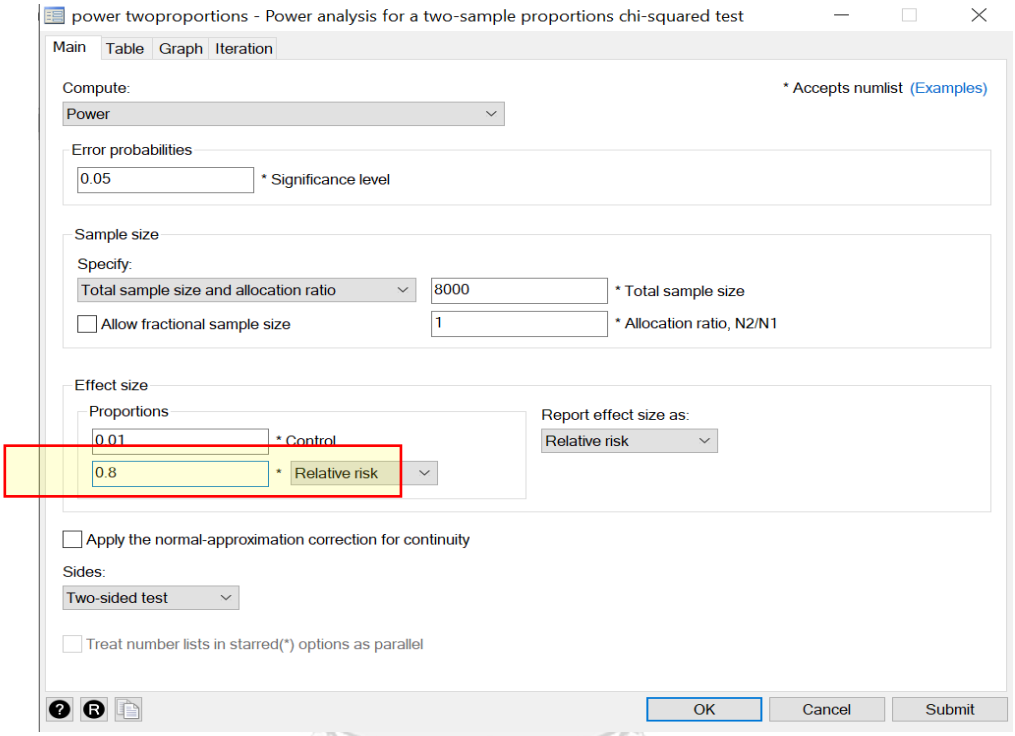

Figure 7 Example of power calculation in STATA program

4.1.4 Similarly, calculate power of all studies included in the meta-analysis.

4.2. Calculate the expected number of studies with statistically significant results (E) by summing the statistical powers from all studies included in the meta-analysis.

SUM of powers = expected number of studies with significant findings (E).

4.3. If E is more than the observed number (O) of studies with significant results, there is no evidence of excess significance based on assumption made for plausible effect size. Findings should be reported as 'not pertinent (NP)'.

4.4. If expected number of studies is less than observed number of studies, the expected number (E) is compared against the observed number (O) of 'positive' studies using the X^2 (chi-square) statistic. Alternatively, one may use a binomial

probability test (preferable with small numbers). Excess statistical significance for single meta-analyses was claimed at P less than .10. The equation used are presented below;

$$A = [(O-E)^2/E+(O-E)^2/(n-E)] \approx X^2$$

Where n = no. of primary studies included in a meta-analysis.