Is the systemic administration of arnica extracts safe? A systematic review of preclinical trials

A administração sistêmica de extratos de arnica é segura? Uma revisão sistemática de ensaios pré-clínicos

¿Es segura la administración sistémica de extractos de árnica? Una revisión sistemática de ensayos pré-clínicos

Abstract:

“Arnica” extracts are widely used in folk medicine to treat acute and chronic inflammatory ailments. Nevertheless, their toxic effects upon systemic use are still not fully understood. Therefore, this work provides a systematic review on the safety of arnica extracts following preclinical trials covering their oral and intraperitoneal administration in animal models. Henceforth, PRISMA guideline was followed and the study protocol was registered in PROSPERO (CDR42020167112). Searches were performed in PubMed (MEDLINE), Scopus, Science Direct, Web of Science (Science Citation Index), and Health Virtual Library (BVS) databases; while SYRCLE’s Risk of Bias tool and CAMARADES checklist were used to assess scientific quality. From 382 articles, five studies met eligibility criteria and underwent qualitative analysis. Data of acute toxicity was reported in all the selected articles, and the treatment time was up to 14 days. Moreover, the following species were reported: Solidago chilensis (hazard categories of 4 and 5 for i.p. and v.o administration, respectively); Solidago microglossa (hazard category of 3, i.p); Lychnophora trichocarpha (hazard category of ≥ 4, i.p); and Lychnophora pinaster (hazard category of ≥ 4, v.o). The alcoholic extracts showcased greater toxic potential, which increased in a dose-dependent manner after 100 mg/Kg. Concerning organ-specific toxicity, the articles reported hepatotoxicity and nephrotoxicity following histopathological analysis. However, the safety of S. chilensis, L. pinaster, L. trichocarpha, and S. microglossa following systemic administration remains unclear due to the limited experimental quality of the included papers, as well as the lack of reported chronic toxicity, pharmacokinetics, and mutagenicity studies. **Keywords:** Administration, Oral; Arnica; Injections, Intraperitoneal; Systematic review; Toxicity.

Resumo:

Os extratos de “Arnica” são amplamente usados na medicina popular para tratar doenças inflamatórias agudas e crônicas. No entanto, seus efeitos tóxicos mediante uso sistêmico ainda não são totalmente compreendidos. Portanto, este trabalho fornece uma revisão sistemática sobre a segurança de extratos de arnica em ensaios pré-clínicos cobrindo sua administração oral e intraperitoneal em modelos animais. Para tanto, a diretriz PRISMA foi seguida e o protocolo do estudo foi registrado no PROSPERO (CDR42020167112). As pesquisas foram realizadas nas bases de dados PubMed (MEDLINE), Scopus, Science Direct, Web of Science (Science Citation Index) e Biblioteca Virtual da Saúde (BVS); enquanto a ferramenta de risco de viés do SYRCLE e o checklist CAMARADES foram usadas para avaliar a
la seguridad de los extractos de árnica después de ensayos experimental de los artículos incluidos, así como a la falta de relatos sobre toxicidad crónica, farmacocinética y estudios de mutagenicidad.

**Palabras clave:** Administración oral; Árnica; Inyecciones intraperitoneales; Revisión sistemática; Toxicidad.

**Resumen:**
Los extractos de "árnica" se utilizan ampliamente en la medicina popular para tratar dolencias inflamatorias agudas y crónicas. Sin embargo, sus efectos tóxicos sobre el uso sistémico aún no se comprenden completamente. Por lo tanto, este trabajo proporciona una revisión sistemática sobre la seguridad de los extractos de árnica después de ensayos preclínicos que cubren su administración oral e intraperitoneal en modelos animales. En adelante se siguió la guía PRISMA y se registró el protocolo del estudio en PROSPERO (CDR42020167112). Las búsquedas se realizaron en las bases de datos PubMed (MEDLINE), Scopus, Science Direct, Web of Science (Science Citation Index) y Health Virtual Library (BVS); mientras que la herramienta Riesgo de sesgo de SYRCLE y la lista de verificación CAMARADES se utilizaron para evaluar la calidad científica. De 382 artículos, cinco estudios cumplieron los criterios de elegibilidad y se sometieron a análisis cualitativos. Los datos de toxicidad aguda se informaron en todos los artículos seleccionados y el tiempo de tratamiento fue de hasta 14 días. Además, se notificaron las siguientes especies: *Solidago chilensis* (categorías de peligro 4 y 5 para administración i.v. respectivamente); *Solidago microglossa* (categoría de perigro 3, i.p); *Lychnophora trichocarpha* (categoría de peligro 4, i.p); e *Lychnophora pinaster* (categoría de peligro 4, v.o). Los extractos alcoólicos presentaron un mayor potencial tóxico, que aumentó de forma dose-dependente después de 100 mg/Kg. Con relación a toxicidad órgano-específico, los artículos relataron hepatotoxicidad e nefrotoxicidad después de análisis histopatológico. No entanto, a seguridad de *S. chilensis*, *L. pinaster*, *L. trichocarpha* y *S. microglossa* después a administración sistémica permanece incerta debido a la calidad experimental limitada de los artículos incluidos, bien como a falta de relatos sobre toxicidad crónica, farmacocinética y estudios de mutagenicidad.

**Palabras clave:** Administración Oral; Árnica; Inyecciones intraperitoneales; Revisión sistemática; Toxicidad.

1. **Introducción**

El género *Arnica* es conocido en el Occidente occidental el más importante granos de especies de interés terapéutico en la medicina popular. La familia de las *Asteraceae* es considerada uno de los más importantes de especies de interés terapéutico en la medicina popular (Dolisgan et al., 2020; Jayasundera et al., 2021; Michel et al., 2020; Panda et al., 2019; Yousuf et al., 2019). Popularmente conocido como daisy, aster, compost o sunflower family, the *Asteraceae* accounts for approximately 1600 genera, 23,000 species and 13 subfamilies (Kenny et al., 2014; Obón et al., 2012). Amongst the species of remarkable therapeutic interest is *Arnica montana* L., also known as Leopard’s Bane. This plant was originally named “arnica” in folk usage, though this name is generically used nowadays for 24 *Asteraceae* species. Nonetheless, all arnicas have been integrated into the folk medicine of many populations throughout the centuries against a range of inflammatory ailments (Kriplani et al., 2017).

In regards to the endemic and exotic herbs used in Brazilian folk medicine and complementary therapeutic practices, several species of “arnica” are noteworthy (Athayde et al., 2019; de Athayde et al., 2021; Marisco et al., 2017; Ribeiro et al., 2017), which are typically used as remedies in the form of tisanes; tinctures; bottled; globules; hot compresses; ointments; and gels. This range of formulations is accompanied by the therapeutic versatility of arnica, which is reported to treat: post-operative lesions such as bruises, sprains and abrasions; edema related to fracture and rheumatic pain of muscles and joints; inflammatory processes of the oropharynx; furunculosis; insect bites and stings; superficial phlebitis, etc (Athayde et al., 2019; Carvalho et al., 2018; Rodríguez-Chávez et al., 2017; Saraiva et al., 2015; Siqueira et al., 2018; Souza et al., 2017).

Amongst the often reported species in ethnopharmacological studies is the Leopard’s Bane (Obón et al., 2012), Mexican arnica (*Heterotheca inuloides*) (Rodríguez-Chávez et al., 2017), and the several “Brazilian arnicas” i.e., *Arnica*
Angustifólia, Arnica chamissonis, Calea uniflora, Chaptalia nutans, Lychnophora pinastere, Lychnophora brasiliensis, Lychnophora ericoideas, Lychnophora diamantinana, Lychnophora saliciifolia, Lychnophora trichocarpha, Solidago chilensis, Solidago microglossia, Porophyllum ruderale, Pseudobrickellia brasiliensis, and Sphagneticola trilobata (Athayde et al., 2019; Baatsch et al., 2017; Bolson et al., 2015; de Athayde et al., 2021; Magalhães et al., 2019; Saraiva et al., 2015; Siqueira et al., 2018; Souza et al., 2017).

Moreover, the therapeutic effects of these species have been demonstrated through pre-clinical studies (Bernardes et al., 2021; Caldeira et al., 2019; da Silva Prade et al., 2020; de Barros et al., 2016; de Sá Müller et al., 2019; Valverde et al., 2020; Vagas et al., 2020) and clinical trials (Iannitti et al., 2016; Mawardi et al., 2020; Raghibi et al., 2018; Silva et al., 2015; Sorrentino et al., 2017), and was attributed to appreciable amounts of several phytochemical markers in the phytocomplex, such as: sesquiterpene lactones; flavonoid glycosides; chlorogenic acid derivatives; tannins; resins; coumarins; carotenoids; inulin; arnicacin; phytosterine; and triterpenes (arnidiol, pradiol and amisterine) (Athayde et al., 2019; Caldeira et al., 2019; de Barros et al., 2016; Meinhart et al., 2017).

Owing to the wide and consecrated therapeutic value of arnica species, there are currently four registered herbal medicines in the Brazilian National Health Surveillance Agency (Anvisa), three based on A. montana (topical use as a healer), and one based on S. chilensis (oral and topical use as anti-inflammatory) (Carvalho et al., 2018). Furthermore, S. microglossia has been included in the List of Medicinal Plants of Interest to the Brazil’s Unified Health System – RENISUS (BRASIL, 2009). However, the safety of the systemic use of arnica tisanes and tinctures has been questioned by both the academic and medical communities (Hudson et al., 2018), especially due to reports of poisoning or intoxication (Bertin, 1864; Canders et al., 2014; Colombo et al., 2010).

As with many other species of medicinal herbs, some arnicas are over-the-counter products. This availability often leads to their use in self-medication without proper professional guidance (Saraiva et al., 2015; Siqueira et al., 2018; Souza et al., 2017). This unsupervised use may become a serious public health problem, when taking into account the toxic potential of their phytochemicals (Hudson et al., 2018). In this sense, the enforcement of the rational use of herbal medicines is of utmost importance, and should be based on the knowledge of their effectiveness and potential risks to the health, as well as consider the reproducibility and constancy of their quality; which must be validated through ethnopharmacological surveys, technical-scientific documentation and/or clinical evidences (Carvalho et al., 2018; Dutra et al., 2016).

Therefore, it is essential that the academic community drive efforts to produce high-level evidences on the safety of these herbal products to ratify their ethnopharmacological claims. In this context, systematic literature reviews regarding arnicas uses in healthcare are noteworthy outreaches to keep healthcare professionals and scientists up to date, as well as provide a starting point for the development of clinical management strategies, which can guide the definition of the limits of what is known and unknown regarding these herbal medicines, and help to formulate hypotheses for further investigations (Colalto, 2018). Owing to the relevance of studies of this nature, we henceforth report the systematic review of the toxicity of arnica extracts in animal models following systemic administration.

2. Methodology

2.1 Study design

This systematic review was performed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline (Moher et al., 2009). Herein, we have followed the systematic review protocol for animal intervention studies presented elsewhere (de Vries et al., 2015). The study followed the registered protocol number PROSPERO CRD42020167112 (Couto et al., 2020).
2.2 Information sources and search strategies

The electronic databases PubMed (MEDLINE), Scopus, Science Direct, Web of Science (Science Citation Index), and Health Virtual Library (BVS) were searched for relevant research papers published up to January 31st, 2020. Thereby, all searches were performed in February 1st, 2020. The definition of the descriptors was performed using the Medical Subject Headings (MeSH) and the Health Sciences Descriptors (DeCS). The search strategy was defined following the PECO mnemonic, i.e., population (P); exposure (E); comparator (C); and outcomes (O). The search terms were combined using boolean logic terms ‘AND’ and ‘OR’. The search strategies used in each database are depicted in Table S1 of supplementary material.

We also evaluated the references of the included studies that have the potential to meet the criteria established in the review. The search was performed by the name or Digital Object Identifier System (DOI) of the intended study through Google. Nor language neither starting date of the collection were restricted during the identification of the studies. The authors of the unavailable articles were contacted twice by e-mail, through which access to these articles and further information were requested.

Search results were extracted and exported to the Rayyan QCRI platform (Ouzzani et al., 2016), which enabled handling duplicates, screening titles and abstracts, performing full text screen, as well as data extraction. Two independent reviewers (VMO and ROC) have identified and removed repeated articles, and carried out the title and abstract screening in duplicate to bypass bias in the selection and exclusion of the papers.

2.3 Study selection

Eligible articles were read in full and assessed regarding predefined inclusion and exclusion criteria, which were prioritized in terms of i) population type, ii) study design, iii) drug type, and iv) publication type, as follows:

**Inclusion criteria** – i) animals of all sex, all age, and all species/strains; ii) experimental models of oral and intraperitoneal intoxication in animals treated with either vehicle/placebo, or health animal control; iii) treatment with arnica preparation (monoherbal) in any dosing, given at any time and frequency of dosing; and iv) original articles and short communications (published or ahead of print).

**Exclusion criteria** – i) studies in humans or animals with any co-morbidity; ii) in vitro, ex vivo, and in silico study designs, before-after studies without control group, observational studies; iii) animal models of oral and intraperitoneal intoxication treated with any other drug/ or treatment with polyherbal preparation of arnica/ or isolated pure compounds/ or arnica combined with standard oral toxicant agents; iv) case reports, review articles, editorials, and letters to the editor, as well as papers presented in scientific events, news, comments, dissertations and theses.

Throughout the study, each reviewer has been blinded to the decisions of the other, and any discrepancies were discussed until consensus was achieved, or, if this was not possible, a third researcher was consulted. It was of utmost importance that the final decision of the articles to be included was taken by consensus between all researches herein involved.

2.4 Data extraction

The data of the included papers was extracted into a pre-piloted standardized data extraction form (electronic spreadsheet). We have collected data in regards to animal model (species/strains of used animals; sex of the animals; number of animals per group; age and weight of animals); study design (number of experimental groups and duration of follow up; method of allocation to treatment group; and method of assessment of oral toxicity whether blinded or not); features of intervention (taxonomical identification of herb; voucher number; method of preparation of arnica extract; quality control methods and parameters; chemical composition; dose; dosing frequency; time and route of administration whether oral or intraperitoneal); study identification (authors; publication year; country where the research was performed); and language of publication).
The primary assessed outcomes were: LD$_{50}$; LD$_{100}$; survival and signs of intoxication or adverse effects (e.g., general behavior). The secondary outcomes were: absolute body weight; results of organ analysis (weight measurement and histopathological study); hematological analysis of red blood cells count (RBC); hemoglobin (Hb); hematocrit (Hct); mean corpuscular volume (MCV); mean corpuscular hemoglobin (MCH); mean corpuscular hemoglobin concentration (MCHC); platelets number; white blood cells count (WBC) and differential leukocyte count; biochemical analysis of alkaline phosphatase (ALP); alanine aminotransferase (ALT); aspartate aminotransferase (AST) and total proteins (PT) for hepatic function; urea nitrogen (BUN); uric acid (AUR) and creatinine (CRE) for renal function; amylase (AMY) for pancreatic function; and results from locomotion and muscle tone evaluation (number of squares covered by the animals, number of rearing, and muscle strength of the animals).

2.5 Risk of bias and quality assessment of the included papers

Two reviewers (VMO, ROC) independently evaluated the risk of bias by using the Systematic Review Center for Laboratory animal Experimentation (SYRCLE’s) tool (Hooijmans et al., 2014). This tool enabled us to evaluate: i) selection bias; ii) performance bias; iii) detection bias; iv) attrition bias; v) reporting bias; and iv) other types of bias. The ten domains in the assessment involved: random sequence generation; baseline characteristics; allocation concealment; random housing; blinding of participants and personnel; random outcome assessment; incomplete outcome data; selective outcome reporting; and other sources of bias. We considered the methodological domains of individual studies as low risk (L), high risk (H), or unclear risk (?) of bias.

Afterwards, two independent investigators (VMO, ROC) performed quality assessment of all included studies by means of the Collaborative Approach to Meta-Analysis and Review of Animal Data from Experimental Studies (CAMARADES) checklist for study quality (Auboire et al., 2018; Macleod et al., 2004). The checklist comprises ten “Yes or No” issues concerning: publication in peer reviewed journal; statement of the temperature control; randomization of treatment or control; allocation concealment; blinded assessment of outcome; avoidance of anesthetic with marked intrinsic properties; use of animals with hypertension or diabetes; sample size calculation; statement of compliance with regulatory requirements; and statement regarding possible conflict of interest. Each item receives a score of “Yes” on the checklist according to their disclosure in the respective article. In the end, “Yes” scores were collected and the total “Yes” score for each paper was given. In total, each article can score up to 10 points, being the average of scores calculated for each paper. Disagreements were resolved by discussion.

2.6 Data synthesis and statistical analysis

Data from included studies was described in narrative synthesis, and summarized in tables and figures to establish patterns and variations. We used signs to indicate statistically significant increase (↑) and decrease (↓) or equal (↔) effect size measured at follow up between treatment and control groups. The Kappa coefficient (Landis & Koch, 1977) was determined to assess the degree of agreement between the two evaluators (VMO, ROC). In this pursuit, we considered a 95% confidence interval and used the Stata 11.0 software package (StataCorp LLC, Texas, USA).

3. Results

3.1 Results of the search

The Figure 1 presents the PRISMA flow diagram with the selection process of the studies in this systematic review. A total of 382 were identified through electronic and manual searching. After removing duplications (n = 18) and screening the articles based on the titles and abstracts, eleven articles remained. These articles were read in full and analyzed according to the
eligibility criteria. Therefore, six articles were excluded due to following reasons: three articles included animals with comorbidity (Rodrigues et al., 2016; Schneider et al., 2015; Sharma et al., 2016); two article did not configure models of oral or intraperitoneal intoxication (Khuda-bukhsh & Chakrabart, 1998; Malpezzi-Marinho et al., 2019); and one research did not evaluate an arnica species (Mishra et al., 2011). Therefore, five studies (Bucciarelli et al., 2010; Facury Neto et al., 2004; Ferrari et al., 2012; Ferreira et al., 2014; Paula-Freire et al., 2014) were included in qualitative synthesis. Throughout the eligibility evaluation, the degree of agreement between the two researchers was considered almost perfect once the Kappa coefficient was of 0.992 (Landis & Koch, 1977).

Figure 1 – PRISMA flow diagram of screened, included and excluded studies in the systematic review of the toxic potential of arnica extracts following systemic administration in animal models

3.2 Description of included studies

The Table 1 summarizes the experimental design of the studies included in the systematic review of the toxicity of arnica’s extracts following systemic administration in animal models. All papers were published in the last two decades and in English language. Also, all publications came from South American countries, being one from Argentina (Bucciarelli et al., 2010) and the great majority (80 %) from Brazil (Facury Neto et al., 2004; Ferrari et al., 2012; Ferreira et al., 2014; Paula-Freire et al., 2014). Overall, four different species of arnicas belonging to two genera have been identified i.e., *S. chilensis* (40 %) (Bucciarelli et al., 2010; Paula-Freire et al., 2014), *S. microglossa* (20 %) (Facury Neto et al., 2004), *L. trichocarpha* (20 %) (Ferrari et al., 2012), and *L. pinaster* (20 %) (Ferreira et al., 2014). It is noteworthy that all studies presented full and accurate information about plant name, identification of specimens, and voucher of the specimen deposited at the herbarium.
Concerning the strains of animal species, the Swiss albino mice (Ferrari et al., 2012; Paula-Freire et al., 2014) were the most commonly used in the experimental models. Swiss mice (Ferreira et al., 2014), Wistar rats (Facury Neto et al., 2004) and CF-1 albino mice (Bucciarelli et al., 2010) have been reported as well. In addition, the majority of the experimental models used animals of both genders (Ferrari et al., 2012; Ferreira et al., 2014; Paula-Freire et al., 2014), while 20% used only male animals (Facury Neto et al., 2004), and 20% used only female animals (Bucciarelli et al., 2010). In the case of using females, none of the studies reported pregnancy, thus fulfilling the requirements of widely cited international guidelines (OECD/OCDE, 2001, 2008).

The age of the used animals was reported in 40% of the studies (Bucciarelli et al., 2010; Paula-Freire et al., 2014). This depended on the animal strain and ranged from 10 up to 12 weeks, which confirms the use of young adult animals. Similarly, the weight of the animals depended on the strain, and two studies (Bucciarelli et al., 2010; Paula-Freire et al., 2014) did not report such information.

As regards the treatment groups (TG), the ethanolic extracts of the aerial parts (EEAP), i.e., flowers, leaves and steams, were the most used arnica derivatives in the studies (Ferrari et al., 2012; Ferreira et al., 2014; Malpezzi-Marinho et al., 2019; Paula-Freire et al., 2014), thereby accounting for 60% of the reports. Conversely, the aqueous extracts of aerial parts (AEAP) accounted for 40% of the evaluated herbal medicines (Bucciarelli et al., 2010; Facury Neto et al., 2004).

In exception of part of one study (Facury Neto et al., 2004), the comparison groups (CG) have been reported accordingly, thus configuring controlled preclinical trials. In this case, the experimental protocol used for assessing the LD$_{50}$ and LD$_{100}$ did not require a comparison (Lorke, 1983; Zbinden & Flury-Roversi, 1981). The CGs (sham treatments) were saline solution (Facury Neto et al., 2004), distilled water (Paula-Freire et al., 2014), standard diet (Bucciarelli et al., 2010), and aqueous surfactant solutions (Ferrari et al., 2012; Ferreira et al., 2014). In a single study the effects of a TG were compared to an untreated control group (Facury Neto et al., 2004).

The intraperitoneal (i.p.) administration route was used in most of the studies (60%), and the oral (v.o.) administration was conducted in 40% of the trials. The researchers opted for administering the various doses of extracts once in 80% of the reports (Bucciarelli et al., 2010; Ferrari et al., 2012; Ferreira et al., 2014; Paula-Freire et al., 2014), and in just one trial the different doses of the extract were administered daily, i.e., repeated dose regimen (Facury Neto et al., 2004). The doses of TGs ranged from 6.3 up to 2,000 mg/Kg. The most frequent dose was 500 mg/Kg, followed by 250 mg/Kg. It was relevant that both the multiple and single-dose regimens were carried out in order to assess the various levels of toxicity in the different physiological compartments (Facury Neto et al., 2004; Ferrari et al., 2012). Only one study (Bucciarelli et al., 2010) performed a single-dose assay (limit test), in which a dose at the upper limit of testing (2,000 mg/Kg) was evaluated (OECD/OCDE, 2001).

Regardless of the experimental design and outcomes, all studies were of acute toxicity in the testing models. Additionally, such studies displayed short-term follow-ups, lasting up to 14 days depending on the nature of the outcomes being assessed. Both the number of experimental groups and the number of animals per group varied widely as a function of the experimental protocols carried out for toxicity assessment. According to the information provided by the researchers, majority (60%) of the studies followed official guidelines (Brasil, 2004; OECD/OCDE, 2001) in the investigation of the toxicity of herbal products (Bucciarelli et al., 2010; Ferrari et al., 2012; S A Ferreira et al., 2014), while others (Facury Neto et al., 2004; Paula-Freire et al., 2014) performed the experiments according to other protocols (Lorke, 1983; Zbinden & Flury-Roversi, 1981).
Table 1 – Summary of the experimental designs of the studies assessing the toxic potential of arnica’s extracts following systemic administration in animal models.

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Country</th>
<th>Plant specie / Voucher number</th>
<th>Animal Strain</th>
<th>Sex</th>
<th>Age / weight</th>
<th>N° of groups</th>
<th>N° of animal per group</th>
<th>Treatment / Comparator</th>
<th>Doses</th>
<th>Administration routes / Frequencies</th>
<th>Intoxication model / Treatment follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facury Neto et al., (2004)</td>
<td>Brazil</td>
<td>Solidago microglossa DC. / 13,495</td>
<td>Wistar rats</td>
<td>♂️</td>
<td>NI / 280 – 360 g, NI / 284 – 340 g</td>
<td>11</td>
<td>6</td>
<td>AEAP / NA, AEAP / Control (untreated) or vehicle (0.9% NaCl)</td>
<td>6.3, 15.6, 18.8, 21.9, 25.0, 28.1, 31.3, 37.5, 46.9, 62.5, and 93.8 mg/Kg, 16.1 mg/Kg</td>
<td>i.p. / daily, i.p. / daily</td>
<td>Acute / 14 days, Acute / 7 and 14 days</td>
</tr>
<tr>
<td>Bucciareli et al., (2010)</td>
<td>Argentina</td>
<td>Solidago chilensis Meyen / AB9</td>
<td>CF-1 albino mice</td>
<td>♀</td>
<td>10 weeks / NI</td>
<td>2</td>
<td>8</td>
<td>AEAP incorporated into the diet / standard diet</td>
<td>2000 mg/Kg</td>
<td>v.o. / once</td>
<td>Acute / 1 and 14 days</td>
</tr>
<tr>
<td>Ferrari et al., (2012)</td>
<td>Brazil</td>
<td>Lychnophora trichocarpha (Spreng.) Spreng. / 20,635</td>
<td>Swiss albino mice</td>
<td>♂️ / ♀</td>
<td>NI / 25-30 g</td>
<td>4</td>
<td>12</td>
<td>EEAP / vehicle (DMSO:Tween:H₂O 1:1:8)</td>
<td>250, 500, and 750 mg/Kg, 1,500 mg/Kg</td>
<td>i.p. / once, i.p. / once</td>
<td>Acute / 0, 1 and 4h; 1, 7 and 14 days, Acute /14 days</td>
</tr>
<tr>
<td>Paula-Freire et al., (2014)</td>
<td>Brazil</td>
<td>Solidago chilensis Meyen / SP 397.047</td>
<td>Swiss albino mice</td>
<td>♂️ / ♀</td>
<td>3 months / NI</td>
<td>7</td>
<td>8</td>
<td>EEAP / vehicle (0.3 mL of distilled H₂O)</td>
<td>30, 100, 300, 500, 750 and 1000 mg/Kg</td>
<td>i.p. / once</td>
<td>Acute / 14 days</td>
</tr>
<tr>
<td>Ferreira et al., (2014)</td>
<td>Brazil</td>
<td>Lychnophora pinaster Mart. / 19,520</td>
<td>Swiss mice</td>
<td>♂️ / ♀</td>
<td>NI / 30 ± 5 g</td>
<td>4</td>
<td>14</td>
<td>EEAP / vehicle (DMSO:Tween:H₂O 1:1:8)</td>
<td>125, 250, and 500 mg/Kg</td>
<td>v.o. / once</td>
<td>Acute / 0, 1 and 4h; 1, 7 and 14 days</td>
</tr>
</tbody>
</table>

i.p., intraperitoneal; v.o., oral administration; NA, non-applicable; NI, non-informed; ♂️, male; ♀, female; AEAP, aqueous extract from aerial parts; EEAP, ethanolic extract from aerial parts; Source: Articles included in this review.
Source: Authors.
Table 2 – Summary of the preparation methods, yield, and chemical composition of the arnica’s extracts used for assessing the toxic potential following systemic administration in animal models.

<table>
<thead>
<tr>
<th>Study ID / Arnica species</th>
<th>Extraction method</th>
<th>Solvent</th>
<th>Plant: solvent ratio (w/v)</th>
<th>Extraction time</th>
<th>Processing</th>
<th>Extract Yield (% w/w)</th>
<th>Identified compounds / analytical method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facury Neto et al., (2004) / S. microglossa</td>
<td>dynamic maceration</td>
<td>deionized H₂O (RT)</td>
<td>NI</td>
<td>24 h</td>
<td>filtration, centrifugation and lyophilization</td>
<td>NI</td>
<td>NI</td>
</tr>
<tr>
<td>Bucciareli et al., (2010) / S. chilensis</td>
<td>infusion</td>
<td>deionized H₂O (90°C)</td>
<td>1:10</td>
<td>10 min</td>
<td>filtration and rotaevaporation</td>
<td>14.0</td>
<td>flavonoids / Shinoda test phenolic compounds / Ferric chloride reaction</td>
</tr>
<tr>
<td>Ferrari et al., (2012) / L. trichocarpha</td>
<td>static maceration</td>
<td>ethanol (RT)</td>
<td>1:10</td>
<td>2 weeks</td>
<td>rotaevaporation (&lt; 40°C)</td>
<td>19.0</td>
<td>NI</td>
</tr>
<tr>
<td>Paula-Freire et al., (2014) / S. chilensis</td>
<td>static maceration</td>
<td>93% ethanol (RT)</td>
<td>1:10</td>
<td>1 month</td>
<td>filtration, rotaevaporation and lyophilization</td>
<td>NI</td>
<td>s 5-O-E-, 3,4-, and 4,5-di-OE-caffeoylquinic acids and rutin / HPLC–DAD–MS and HPLC–MS/MS</td>
</tr>
<tr>
<td>Ferreira et al., (2014) / L. pinaster</td>
<td>static maceration</td>
<td>ethanol (RT)</td>
<td>NI</td>
<td>2 weeks</td>
<td>rotaevaporation (&lt; 40°C)</td>
<td>12.4</td>
<td>NI</td>
</tr>
</tbody>
</table>

NI, non-informed; GC, gas chromatography; NMR: nuclear magnetic resonance; HPLC, High Performance Liquid Chromatography; DAD, Diode Array Detector; MS, Mass Spectrometer; RT, room temperature; UV, Ultraviolet detector; Source: Articles included in this review.

Source: Authors.
Table 3 – Summary of the primary and secondary outcomes of the studies assessing the toxic potential of arnica extracts following systemic administration in animal models.

<table>
<thead>
<tr>
<th>Study ID / Preparation</th>
<th>Lethal dose / death reports</th>
<th>Biochemical parameters (as compared to control groups)</th>
<th>Physiological and pharmacological parameters</th>
<th>Histopathology, Anatomopathological analysis</th>
<th>Behavioral tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facury Neto et al., (2004) / AEAP <em>S. microglossa</em></td>
<td>LD₅₀: 54.7 ± 0.3 mg/Kg. LD₉₀: 86.2 mg/Kg / No deaths were seen with doses up to 40 mg/Kg</td>
<td>↔ALT ↔AST</td>
<td>ND</td>
<td>The TG exhibited a similar dietary intake as compared to CG. There were no significant differences between TG and CG when analyzed the body weight.</td>
<td>No significant differences were observed between CG and TG in the different parameters analyzed during home cage, hand-held, open field observations, motor activity evaluations, and emotionality parameters (number of groomings and fecal bolus)</td>
</tr>
<tr>
<td>Bucciareli et al., (2010) / AEAP <em>S. chilensis</em></td>
<td>ND / No deaths were observed along the follow-up</td>
<td>ND</td>
<td>ND</td>
<td>The histopathological examinations of brain, liver, kidney, spleen, stomach, and intestine in all animals did not show any changes as consequence of the exposure to TG</td>
<td></td>
</tr>
<tr>
<td>Ferrari et al., (2012) / EEAP <em>L. trichocarpha</em></td>
<td>ND / No deaths regardless of the TG were observed along the follow-up</td>
<td>↔ALT ↔AST ↔ALP ↔PT ↔CRE ↔BUN ↔AUR</td>
<td>ND</td>
<td></td>
<td>Open-field test: The TG (750 mg/Kg) presented locomotion reduction after 1h and 4h of administration as compared to CG; The TGs (500 and 750 mg/Kg) presented reduction of rearing after 1h and 4h of administration as compared to CG; Traction test: The TGs (750 and 250 mg/Kg) presented a reduction on muscle tone after 24h and 14 days of administration, respectively, as compared to CG.</td>
</tr>
<tr>
<td>Paula-Freire et al., (2014) / EEAP <em>S. chilensis</em></td>
<td>LD₅₀: 512.5 mg/kg / 75% of the animals treated with 750 mg/kg and all animals treated with TG at 1000 mg/Kg died within 2 h of administration</td>
<td>ND</td>
<td>No significant change was observed in the body weight; There were no differences in the liver weight of the TGs and CG, but a significant difference was observed in the color; TG (500 mg/Kg) rendered an increase in</td>
<td>The TGs exhibited spots and darkening in the liver in a dose-dependent manner; The TGs presented congestion and dilatation of the sinusoidal space, focal hepatic necrosis, steatohepatitis with evident vacuolar degeneration and</td>
<td>The TG (30 mg/Kg) did not cause any behavioral changes. The TGs (&gt; 100 mg) showed a dose-dependent reduction in general activity; Reduced grip strength, ptosis, cyanosis, piloerection and an increase in the respiratory</td>
</tr>
</tbody>
</table>

*Note: ALT = Alanine aminotransferase; AST = Aspartate aminotransferase; ALP = Alkaline phosphatase; PT = Prothrombin time; CRE = Creatinine; BUN = Blood urea nitrogen; AUR = Auremate; ND = Not determined.*
<table>
<thead>
<tr>
<th>Study</th>
<th>Species</th>
<th>Treatment</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ferreira et al., (2014) / EEAP <em>L. pinaster</em></td>
<td>ND / No deaths regardless of the TG were observed along the follow-up</td>
<td>kidney weight; There was an increase in the weight of spleens from animals treated with TG (300 and 500 mg/Kg); The TGs (300 and 500 mg) significantly decreased food intake as compared to CG</td>
<td>There were hemorrhage signs (depending on the dose). No morphological alterations were observed in the spleens irrespective to the dose</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>No alterations in the lungs and brains of animals from TGs were observed. The kidneys of animals of both genders treated with TGs presented time and dose-dependent alterations such as tubular degeneration without glomerular loss, glomerular congestion and a slight inflammation, tubular vacuolar degeneration, glomerular loss and renal tubules dilation with eosinophilic deposition into the tubules’ lumen, beside moderate inflammatory infiltrate. The hepatotoxicity was evidenced due to vacuolar degeneration, venular congestion, sinusoidal congestion, parenchymal and perivascular inflammatory infiltrate, focal hepatic necrosis, and hydropic degeneration. Although hepatotoxicity was not time and dose-dependent, the TGs injured the livers of females in a greater extent than males.</td>
</tr>
</tbody>
</table>

| Water consumption as well as food intake in all groups were similar. The body weight was not affected in any group. TGs did not induce significant changes to the relative weight of the liver, kidneys, lungs and brain. There were no differences between males and females. |

| Open-field Turner test: The TGs (250 and 500 mg/Kg) showed impaired locomotion without altering the muscle strength after 4 and 24h of administration when compared to CG |

| ↔WBC |
| ↔RCB |
| ↔Hb |
| ↔Hct |
| ↔Platelets |
| ↔MCV |
| ↔MCH |
| ↔MCHC |
| ↔Eosinophils |
| ↔Neutrophils |
| ↔Lymphocytes |
| ↔Monocytes |
| ↔BUN |
| ↔AUR |
| ↔CRE |
| ↔AMY |
| ↓AST* |
| ↓ALT* |
| ↓ALP |
| ↓PT |

*Statistically significant differences at p < 0.05; AEAP, aqueous extract from aerial parts; CG, comparison group; CNS, central nervous system; EEAP, ethanolic extract from aerial parts; ND, non-determined; TG, treatment group. Source: Articles included in this review.

Source: Authors.
3.3 Obtainment and phytochemical characterization of the arnica’s extracts

The Table 2 depicts the summary of the main information regarding the preparation of the arnica’s extracts in the studies herein included, as well as the results of their phytochemical characterization. The static maceration was the most common method (Ferrari et al., 2012; Ferreira et al., 2014; Paula-Freire et al., 2014), even though dynamic maceration (Facury Neto et al., 2004) and infusion (Bucciarelli et al., 2010) were also reported. It is worth of note that we have identified an inconsistence in the description of the extraction method reported by Facury Neto et al. (2004), which stated using infusion technique while in fact the authors have used dynamic maceration.

Water and, mainly, ethanol were the chosen solvents in the preparation of the herbal derivatives. When informed, the drug:solvent ratio was of 1:10 v/w (100 g/L) irrespective to the extraction technique, temperature, and used solvent (Bucciarelli et al., 2010; Ferrari et al., 2012; Paula-Freire et al., 2014). The extraction time varied widely from 10 min up to 1 month depending on the used technique, nevertheless the time was typically greater when dynamic maceration was chosen. In all cases, the extracts were fully concentrated prior to dilution and administration on the animals. Furthermore, the majority of the studies reported the extraction yields, which ranged between 12 and 20 % w/w. Based on the same plant:solvent ratio, the maceration (soaking) of arnicas with ethanol rendered greater extraction of phytochemicals as compared to infusion process. Only two studies (Bucciarelli et al., 2010; Paula-Freire et al., 2014) showcased information regarding the qualitative identification of phytochemicals in the TG, which were carried out with S. chilensis and have reported phenolic compounds such as phenolic acids and flavonoids.

3.4 Systemic acute toxicity of arnica’s extracts

3.4.1 Primary outcomes – LD<sub>50</sub>, LD<sub>100</sub>, death reports and general behavior

The Table 3 summarizes the main primary and secondary outcomes of the preclinical trials assessing the toxicity of arnica’s extracts following systemic administration. Two studies evaluated the LD<sub>50</sub> (Facury Neto et al., 2004; Paula-Freire et al., 2014), and one study evaluated LD<sub>100</sub> for determining the toxicity of the extracts (Facury Neto et al., 2004), all for species of Solidago gender. Both the LD<sub>50</sub> and LD<sub>100</sub> following i.p. administration of aqueous extract from the aerial parts (AEAP) from S. microglossa were between 50 and 300 mg/Kg (Facury Neto et al., 2004), thus presenting a difference of around 58% from each other. Herein, the LD<sub>50</sub> and LD<sub>100</sub> were 3.4 and 5.4 times greater, respectively, than the therapeutic dose known to be effective in wound healing (16.1 mg/Kg), and the i.p. administration of AEAP from S. microglossa has shown to be safe up to 40 mg/Kg.

The LD<sub>50</sub> following i.p. administration of ethanolic extract from the aerial parts (EEAP) from S. chilensis was greater than 300 mg/Kg and lower than 2,000 mg/Kg (Paula-Freire et al., 2014), and the LD<sub>100</sub> was considered 1000 mg/Kg, which is around 95% greater than the respective LD<sub>50</sub>. Neither the v.o. administration of AEAP from S. chilensis (Bucciarelli et al., 2010) and of EEAP from L. pinaster (Ferreira et al., 2014), nor the i.p. administration of EEAP from L. trichocarpha (Ferrari et al., 2012) exhibited any mortality level.

Behavioral analysis was carried out by means of: general observation (Hippocratic screening); motor activity tests (open field, traction test); emotionality parameters (number of groomings and fecal bolus); and dietary (food and water) intake. Other parameters evaluated were absolute body weight, and relative weight of organs.

A study (Bucciarelli et al., 2010) stated that there was not important behavioral changes following the i.p. administration of the AEAP from S. chilensis (2,000 mg/Kg). A similar trend was observed when the EEAP from S. chilensis was administered i.p., but in the lowest dose, i.e., 30 mg/Kg (Paula-Freire et al., 2014). However, when greater doses (100 up to 1000 mg/Kg) were i.p. administered; the locomotion of the animals was remarkably impaired in a dose-dependent manner irrespective of the sex, thereby showcasing a remarkable depressant effect of EEAP on the central nervous system (CNS). Moreover, it was reported an increase in the weights of kidney and spleens, as well as a noticeable decrease in the food intake (Paula-Freire et al., 2014).
All the other studies (Ferrari et al., 2012; Ferreira et al., 2014) did not report any significant alteration in body weight and dietary intake pattern. Concerning the *Lychnophora* genus (Ferrari et al., 2012; Ferreira et al., 2014), regardless the species and administration routes, the EEAP (250 – 750 mg/Kg) also inhibited the motor activity of the animals, which indicates important toxicity to the CNS.

### 3.4.2 Secondary outcomes – biochemical and histopathological analysis

Four studies carried out histopathological analysis on the several physiological compartments of the animals (Bucciarelli et al., 2010; Ferrari et al., 2012; Ferreira et al., 2014; Paula-Freire et al., 2014). A single study (Bucciarelli et al., 2010) indicated that there was no significant changes in histopathological findings following the v.o. administration of the AEAP from *S. chilensis* (2,000 mg/Kg).

Nonetheless, all the other studies (Ferrari et al., 2012; Ferreira et al., 2014; Paula-Freire et al., 2014) reported remarkable signs of hepatotoxicity related to the administration of the EEAP, regardless the species, dose, and administration route. Herein, for *S. chilensis* the hepatotoxicity was dose-dependent (Paula-Freire et al., 2014); while for *L. pinaster* it has shown to be time, dose and sex-dependent (Ferreira et al., 2014). Other histopathological findings revealed toxicity to kidneys, brain, and lungs in a certain extent (Ferrari et al., 2012; Ferreira et al., 2014; Paula-Freire et al., 2014). These last two depended on the species, being more evident to *L. trichocarpa* (Ferrari et al., 2012). Herein, there were no reports of splenic toxicity.

Three studies reported results of hematological and/or biochemical analysis of blood samples withdrawal from the animals (Facury Neto et al., 2004; Ferrari et al., 2012; Ferreira et al., 2014). Only the v.o. administration of EEAP from *L. pinaster* at the doses of 125, 250 and 500 mg/Kg all significantly affected the animals health by decreasing the serum levels of alanine transaminase (ALT) and aspartate transaminase (AST) in relation to CG, which was interpreted as a sign of kidney failure (Ferreira et al., 2014). The other biochemical markers remained unaltered throughout all the other investigations.

### 3.5 Risk of bias and methodological quality

The Figure 2 presents the results of risk of bias assessment according to SYRCLE’s tool for each preclinical trial herein included. Overall, the studies did not present high risk of bias for the evaluated domains, and all studies showcased a low risk of bias in four domains. In addition, unclear risk of bias was verified in the same six domains for all studies, of which two (33.3%) were related to blinding issues. In such cases, we were not able to define whether the risk of bias was low or high due to the lack of accurate information throughout the reports.
Figure 2 - Reporting of risk of bias for the studies assessing the toxic potential of arnica’s extracts following systemic administration in animal.

<table>
<thead>
<tr>
<th>Source: Authors.</th>
</tr>
</thead>
</table>

In turn, the Figure 3 reports the methodological quality as assessed by the CAMARADES checklist for the preclinical trials included. All studies were: published in peer-reviewed journals; avoided the use of anesthetics that could interfere in the evaluated outcomes; have stated to follow the regulatory requirements regarding animal usage in experiments; did not use animals with metabolic disorders; and did not report conflicts of interest. Likewise, the majority of the studies stated to control the environmental conditions during the experimental time frame, and a single study reported randomization for both the TG and CG. Notwithstanding, none of the included studies provided complete information regarding the sample size calculation, allocation concealment, and blinding throughout the assessment of the outcomes; and the methodological quality scores of the studies ranged from five to six with a mean of 5.8.

Figure 3 - Reporting of quality indicators for the studies assessing the toxic potential of arnica’s extracts following systemic administration in animal.

| Source: Authors. |
4. Discussion

Considering that the generic name “arnica” may refer to 24 Asteraceae members, it is noteworthy that all plants reported in the selected articles are considered endemic to Brazilian flora; thereby highlighting the relevance of this region in phytotherapy, ethnopharmacology, and ethnotoxicology. In regards to *S. chilensis* and *S. microglossa*, these species are widespread in the whole country in exception to Amazonia (Halevy, 2018), while either the *L. pinaster* and *L. trichocarpha* are endemic to Brazilian savanna (Athayde et al., 2019); which is the second greatest biome in Brazil and a great source of healing herbs (Ribeiro Neto et al., 2020). Nevertheless, the natural occurrence of these species of medicinal importance is being threatened by the extensive growth of mining; urban sprawl; uncontrolled tourism; cattle raising and criminosus fires (dos Santos et al., 2020).

The therapeutic properties of the several extracts from *S. chilensis* have been extensively revised elsewhere (Assini et al., 2013; Mercandeli et al., 2012). While the AEAP of *S. microglossa* rendered systemic effects upon i.p. administration (16 mg/Kg) in healing open cutaneous wounds in rats (Facury Neto et al., 2004), the EEAP of *S. microglossa* (200 mg/Kg, v.o.) comprising gallic acid and the flavonoids quercitrin (quercetin-3-O-rhamnoside), rutin (quercetin-3-O-rutinoside) and quercetin showcased significant hepatoprotective activity in paracetamol induced liver damage in mice (250 mg/Kg, v.o.) in a dose dependent manner (Sabir et al., 2012). No deaths were reported following the administration of such *S. microglossa* extracts at these doses.

The EEAP from *L. pinaster* (40, 125, 375 mg/Kg, v.o) and their isolated phytochemicals E-lychnophoric, chlorogenic, cinnamic and caffeic acids; rutin; lupeol; and stigmasterol (15 mg/Kg) showcased anti-inflammatory and anti-hyperuricemic activities in mice comparable to allopurinol (10 mg/Kg, v.o.) and indomethacin (3 mg/Kg, v.o.) (de Sá Müller et al., 2019). Furthermore, the EEAP from *L. pinaster* and its fractions of increasing polarity namely hexane, dichloromethane, ethyl acetate, and methanolic (100 mg/Kg, v.o) yielded antinociceptive effects comparable to indomethacin and morphine in experimental models of contortion induced by acetic acid and hot plate, respectively. These effects would be likely correlated with the presence of phytochemicals of medium to high polarity such as terpenes, flavonoids and phenolic acids, which were capable for acting either at peripheral and central levels by controlling the release of substances that induce nociception (Ferreira et al., 2019). In turn, the EEAP of *L. trichocarpha* (25, 125, 250 mg/Kg, v.o) was able to reduce serum urate levels in hyperuricemic mice in a dose-dependent manner, though this outcome was in part attributed to the synergism between the apigenin; luteolin; lupeol; and eremanholide C (25 mg/Kg) (Souza et al., 2012). Henceforth, when considering the findings in these reports, the EEAP of *L. pinaster* and *L. trichocarpha* stand out as natural options for the management of gouty arthritis.

The greater prevalence of investigations which were carried out with ethanolic extracts may be justified by the attempt to mimic the popular use of arnica-based remedies; which are commonly found in bottled preparations. These preparations are very common in Brazilian folk medicine, and comprise the use of one or several medicinal herbs, which are soaked in sugarcane brandy (“cachaça”); white wine; or grain alcohol. Since Brazil is a continental and tropical country, the consumption profile of the different kinds of herbal beverages may vary widely depending on the region (Zank & Hanazaki, 2017). In this context, the consumption of aqueous extracts prepared by infusion or decoction (tisanes) is expected to be much greater in the southernmost states due to the significantly lower average temperature throughout the year. On the other hand, in the warmest regions, the consumption of bottled ethanolic extracts is more evident.

From a phytopharmaceutical technology point of view, the maceration process using an alcoholic solvent may lead to an enhanced extraction efficiency of both the phenolic acids and flavonoids, when compared to boiled water-based methods such as infusion and decoction (Llorent-Martínez et al., 2020; Meinhart et al., 2020). The higher yield of alcoholic extraction may be attributed to the intermediate eluotropic strength of methanol and ethanol (64.7 and 78.5, respectively) when compared to water (100), what therefore enhances the mass transfer of low molecular weight polyphenolic compounds (Stalikas, 2007) such as:
caffeoylquinic acid derivatives; terpenes; glycoside flavonoids; and flavonoids. Nonetheless, these compounds were identified in ethanolic extracts of *S. chilensis* (Paula-Freire et al., 2014), *S. microglossa* (Sabir et al., 2012), *L. trichocarpha* (Souza et al., 2012), and *L. pinaster* (de Sá Müller et al., 2019).

In general, there were reported meaningful histopathological and behavioral alterations following the systemic administration of the arnica’s alcoholic extracts in the animals. In this regard, the hepatotoxicity; depressant-like effect in the CNS; and nephrotoxicity are noteworthy (Ferrari et al., 2012; Paula-Freire et al., 2014), even though the several complementary biochemical analyses did not reveal markedly alterations owing to the nature of the intervention (short-time exposure). Nonetheless, significant changes in the biochemical markers of liver and kidney are known to appear later-on the onset of the injury (Ferrari et al., 2012).

The acute toxicity of the AEAP from *S. microglossa* was around 10-fold greater than that obtained for the EEAP of *S. chilensis*. Both extracts were administered i.p. though in a different regimen (i.e., daily and once, respectively). Based on the results of LD$_{50}$, these remedies belong to the hazard categories 3 (50 mg/Kg < LD$_{50}$ < 300 mg/Kg, toxic if swallowed) and 4 (300 mg/Kg < LD$_{50}$ < 2,000 mg/Kg, harmful if swallowed), respectively. Moreover, the oral administration of the tisane (infusion – AEAP) from *S. chilensis* fulfilled the requirements to be framed in the hazard category number 5 (2,000 mg/Kg < LD$_{50}$ < 5,000 mg/Kg, may be harmful if swallowed). In this category chemicals are acknowledged to be of relatively low acute toxicity but, under certain circumstances, may pose a hazard to especially vulnerable populations (OECD/OCDE, 2001, 2008). Beside the administration route, the solvents used for obtaining the extracts can explain the differences in the toxicity between the “bottled” and the tisanes for both the *S. chilensis* and *S. microglossa*.

Altogether, the findings of biochemical and histopathological analysis suggest that the EEAP from *L. pinaster* have similar or even slightly greater toxicity than that of *L. trichocarpha*. Taking into account the absence of deaths reported following the use of ethanolic preparations for these species in the greater doses (500 and 1500 mg/Kg, respectively), they indirectly should be framed at the most unfavorable scenario in the hazard category number 4.

Even though not representing the real practice of drug administration in the folk medicine, the i.p. route renders very useful data concerning the assessment of toxicity for herbal extracts and other preparations of medical interest, thus being widely used for the pharmacological and toxicological validation of these products. This is because this administration route, when compared to oral, is less harmful for the animals and outreaches a greater and prompt bioavailability of macromolecules by avoiding both the gut flora metabolism and the animal refusal during feeding (Al Shoyaib et al., 2020). In this sense, the LD$_{50}$ of a drug when administered v.o. is considerably higher than i.p. LD$_{50}$.

Regarding the LD$_{50}$ of drug preparations according to the route of administration, ibuprofen was reported to yield in mice values of 800 mg/Kg for v.o. and 320 mg/Kg for i.p (Adams et al., 1969). This over-the-counter non-steroidal anti-inflammatory is widely used in the management of rheumatoid and osteo-arthritis, as well as several inflammatory illnesses (Bushra & Aslam, 2010). Nonetheless, ibuprofen is a category 4 compound, just as the majority of the “Brazilian arnicas” herein discussed. In addition, the psychoactive pseudoalkaloid caffeine, which can be found in coffee and in a broad range of beverages including medicinal teas and tisanes (van Dam et al., 2020) showcased LD$_{50}$ of 367 mg/Kg (v.o) in rats (Adamson, 2016), thereby meeting the hazard category 4 criteria. Notwithstanding, the acute toxicity of caffeine following oral intake by humans is not often discussed in the literature (van Dam et al., 2020).

In the regard of evidence-based safety issues, chronic toxicity tests as well as an investigation of the possible mutagenic effects of arnicas would be of utmost importance to shed light on their overall preclinical safety profiles (Oyedepo & Palai, 2021). On the same hand, a comprehensive pharmacokinetic evaluation; adequate pharmacodynamic assaying; and clarifying the mechanisms of toxicological action and the major adverse effects would be also crucial to attend to regulatory standards and permit the post-registration monitoring in Brazil (Carvalho et al., 2018).
Animal studies are often substandard predictors of human responses to drug exposures (Bracken, 2009; Leenaars et al., 2019), especially with regards of predicting toxicity in clinical trials (Van Norman, 2019, 2020). The lack (i.e., being neither recognized nor reported) or misuse of blinding strategies might be considered the greater source of bias and, therefore, the utmost methodological weakness of the included papers. Together with several other issues, this may significantly impair the extrapolation of animal data to humans because it may culminate in the induction of the expected result by the researcher (Bracken, 2009). Noteworthy, the evaluation of the methodological quality of the included papers is one of the main strengths of a systematic review as compared to the traditional narrative reviews (Colalito, 2018).

Therefore, though previous reports on the therapeutic benefits of several arnica preparations were carried out at lower doses following full concentration of the extracts (Assini et al., 2013; de Sá Müller et al., 2019; Souza et al., 2012; Facury Neto et al., 2004; Malpezzi-Marinho et al., 2019; Mercandeli et al., 2012; Sabir et al., 2012; Schneider et al., 2015), and this concentration does not occur in the real traditional therapeutic practice; it is still not conceivable to guarantee whether the folk claims for the systemic use of “Brazilian arnicas” in acute inflammatory ailments is safe or not.

To the best of our knowledge there are not case reports or other scientific evidences of human acute toxicity following the systemic administration of herbal preparations based in “Brazilian arnicas”. The Brazilian sanitary legislation enables the register of herbal remedies as traditional herbal products as long as their safety have been showcased based on traditional usage (Carvalho et al., 2018). Thereby, even being relatively more toxic in preclinical trials as compared to the other arnica species herein evaluated, the oral intake of S. microglossa for treating inflammatory ailments has been recently quoted in the List of Over-the-Counter Drugs by the Anvisa (Brasil, 2021).

Even fulfilling all the safety assumptions aforementioned, should “Brazilian arnica’s” preparations be integrated into healthcare by health providers at the basic assistance level, the general condition of the patients must be fully understood prior prescribing these phytomedicines by oral route. In this sense, it would be recommendable to request regular exams of renal and hepatic function in order to follow the evolution of the treatment and mitigate adverse effects. Dealing with the feasibility of dose extrapolation from animals to humans, normalization methods based on body surface area should be preferred rather than simple conversions based purely on body weight(Shin et al., 2010).

Additionally, the indiscriminate and unadvised exposure to arnicas extracts by systemic administration at a chronic basis may pose unknown risks for the health of the patients, and considering that inflammatory conditions may last long according to their severity, the use of these herbal medicines might only be advisable upon full validation of safety and efficacy in long-term use. The lack of chronic toxicity outcomes; the reduced number of included studies, beside their limited methodological quality; and the impossibility of quantitatively comparing their outcomes by mean a meta-analysis owing to the great difference in the experimental variables are the main limitations of this study.

5. Conclusions

This unprecedent systematic review gathered evidences of preclinical trials that made it possible to understand that it is still not reasonable to categorically assert that the systemic administration of S. chilensis, L. pinaster, L. trichocarpha, and S. microglossa is safe; at least until the contrary be fully proved by high quality experimental models, clinical trials and pharmacovigilance studies. Owing to the remarkable hepatotoxicity and nephrotoxicity of these “Brazilian arnicas” in animal models, their use should be therefore contraindicated to patients suffering from hepatic and/or renal impairment. Furthermore, this work: i) highlights that the use of phytomedicines by Brazilian folks should be followed up by qualified and updated healthcare professionals; ii) corroborates the urge of revisioning the common sense that “natural products don’t hurt at all”, and iii) reinforces the need of enhancing the methodological quality of preclinical studies for enabling credible animal to human translation.
Further studies devoted to elucidate the pharmacokinetics, pharmacodynamics and possible adverse effects of the oral administration of arnica’s extracts are required to ensure their safe use in humans. In upcoming investigations, our research group will shed light to the evidences concerning the systemic toxicity of other herbs commonly used in the complementary and alternative medicine in Brazil.

Conflicts of interest statement

None.

Acknowledgments

The authors gratefully thank for the financial support proved by the Universidade Federal de São João del Rei and Minas Gerais Research Funding Foundation (PIBIC UFSJ/FAPEMIG grant 19521). This study was also financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) - Finance Code 001.

Author’s contributions

VMO acquired, treated and analyzed data; as well as wrote the first draft of the manuscript. FMDC and NRB aided in the conceptualization of the study and revised the manuscript. DVT revised data and writing. ROC conceptualized the research; as well as coordinated all data collection, treatment, analysis and wrote the final version of the manuscript.

References


