

Use and Safety of Anthroposophic Medicinal Products: An Analysis of 44,662 Patients from the EvaMed Pharmacovigilance Network

Harald Johan Hamre^{1,2}  · Anja Glockmann^{1,2}  · Kirsten Heckenbach³ · Harald Matthes³ 

Published online: 30 September 2017
© The Author(s) 2017. This article is an open access publication

Abstract

Background There is a need for data on the clinical safety of anthroposophic medicinal products (AMPs).

Objectives The main objective of this analysis was to determine the frequency of adverse drug reactions (ADRs) to AMPs, relative to the number of AMP prescriptions.

Methods EvaMed (Evaluation of Anthroposophic Medicine) was a prospective pharmacovigilance study with the patients of 38 physicians in outpatient care in Germany. Diagnoses and prescriptions were extracted from the electronic medical records. All physicians documented ADRs of Grades III–IV and serious ADRs, seven ‘prescriber physicians’ also documented non-serious ADRs of any intensity. Patients were eligible for this analysis if they had one or more AMP prescription in the years 2001–2010, followed by one or more physician visit.

Results A total of 44,662 patients with 311,731 AMP prescriptions, comprising 1722 different AMPs, were included. One hundred ADRs to AMPs occurred, caused by 83 different AMPs. ADR intensity was mild, moderate, and severe in 50% ($n = 50/100$), 43%, and 7% of cases, respectively; one ADR was serious. Among patients of prescriber physicians, ADRs of any intensity occurred in 0.071% ($n = 67/94,734$) of AMP prescriptions and in 0.502% ($n = 65/12,956$) of patients prescribed AMPs. In subgroup analyses according to age, specific AMPs or AMP groups, dosage forms, and concentrations (altogether 11 groups), the highest ADR frequency was 0.290% of prescriptions (for one specific AMP). Among all patients, serious ADRs occurred in 0.0003% ($n = 1/311,731$) of prescriptions and 0.0022% ($n = 1/44,662$) of patients.

Conclusion In this analysis from a large sample, ADRs to AMP therapy in outpatient care were rare; ADRs of high intensity as well as serious ADRs were very rare.

Electronic supplementary material The online version of this article (doi:10.1007/s40801-017-0118-5) contains supplementary material, which is available to authorized users.

✉ Harald Johan Hamre
harald.hamre@ifaemm.de

¹ Institute for Applied Epistemology and Medical Methodology, University of Witten/Herdecke, Zechenweg 6, 79111 Freiburg, Germany

² European Scientific Cooperative on Anthroposophic Medicinal Products, Zechenweg 6, 79111 Freiburg, Germany

³ Research Institute Havelhöhe, Community Hospital Havelhöhe, Kladower Damm 221, 14089 Berlin, Germany

Key Points

This safety analysis comprised 44,662 outpatients with altogether 311,731 prescriptions of anthroposophic medications.

Adverse reactions to these medications were rare (0.071% of prescriptions), serious adverse reactions were very rare (0.0003%).

In this analysis, anthroposophic medication therapy was a safe treatment.

1 Background

1.1 Pharmacovigilance of Medicinal Products Used in Whole Medical Systems

The clinical safety of drugs (medicinal products, MPs) on the market is an ongoing public health issue and of paramount importance for drug regulation. Pharmacovigilance of MPs is defined as the science and activities relating to the detection, assessment, understanding, and prevention of adverse effects or any other drug-related problems. Methods used in pharmacovigilance include pre-clinical testing, spontaneous reporting of suspected adverse drug reactions (ADRs), case-control studies, cohort studies, database analyses, and clinical trials [9]. National and supranational pharmacovigilance centers (e.g., Uppsala Monitoring Centre of the WHO [8], European Medicines Agency [3]) are central players in initiating and coordinating pharmacovigilance activities, but initiatives may also come from other stakeholders.

Modern pharmacovigilance has become differentiated to accommodate special product groups, including herbal [11] and other MPs used in whole medical systems (WMPs) [7, 13, 43]. Of these WMPs, some (e.g., those used traditionally in China [43] and on the Indian subcontinent [13]) were historically produced for local use. In modern times, industrial scale production developed without the rigorous quality control of modern drug manufacturing, and in many countries these WMPs are regulated as food or food supplements or are imported for use without any regulation. Some of these WMPs have been associated with repeated, severe ADRs including toxicity, and current pharmacovigilance initiatives towards these WMPs has to deal with specific challenges (e.g., classification issues, lack of standardization, contamination by heavy metals or pesticides [12]).

For the WMPs used in homeopathy and anthroposophic medicine (AM), the historical and regulatory situation is different: since the last quarter of the 20th century, they have been marketed in European countries such as Austria, France, Germany, and Switzerland as drugs, manufactured according to Good Manufacturing Practice standards, and subject to modern drug regulation. Toxicologically relevant starting materials (e.g., aconite, cinnabar) are highly diluted according to the safety requirements of European regulations [7, 15, 32]. Adverse reactions to these WMPs are infrequent and usually of mild to moderate severity; anaphylactic reactions occur but are very rare [14, 27, 30]. However, the large majority of these WMPs were introduced on the European market before clinical trials became widespread, and for many MPs safety data from clinical studies are still sparse. Also, there is a need for clinical safety data on children.

A feature of all WMPs is the large number of different MPs used within each medical system, often alongside non-medication therapies. To accommodate for this and other features outlined above, a five-step research strategy for whole medical systems and their WMPs has been proposed: (1) context, paradigms, philosophical understanding, and utilization; (2) safety status; (3) comparative effectiveness; (4) component efficacy; (5) biological mechanisms [20]. In this model, effect evaluation takes place first on the level of the whole system (Step 3: comparative effectiveness) and subsequently on the level of system components (e.g., individual WMPs). Similarly, safety assessment (Step 2) includes analyses on the level of the whole medical system as well as for individual system components [29]. In this context, the present analysis concerns MPs used in the whole medical system of AM (AMPs), and its primary scope is use (Step 1) and safety (Step 2) on the level of the whole system.

1.2 Anthroposophic Medicinal Products and the EvaMed Study

AM was developed in the early 1920s in Germany and Switzerland and is now practiced by an estimated 19,000 physicians around the world [4]. AM therapy is employed in inpatient and outpatient settings in most medical fields and involves a large number of different AMPs, with more than 1300 AMPs marketed in Germany in the years 2012–2014. AMPs are manufactured from plants, minerals, animals, and from chemically definable substances [26]; quality standards for starting materials and manufacturing procedures are described in the European Pharmacopoeia (Ph.Eur.), in German [Deutsches Arzneibuch (DAB), Deutsches Homöopathisches Arzneibuch (HAB)], French (Ph.fr.) or Swiss (Ph.Helv.) pharmacopoeias or in the Anthroposophic Pharmaceutical Codex (APC) [26]. Manufacturing procedures for AMPs include standard procedures used for the manufacturing of homeopathic or herbal MPs as well as special procedures such as the production of metal mirrors (deposits of metals in reduced state onto a surface [26]) by chemical vapor decomposition; the processing of herbs by fermentation, toasting, carbonizing, incineration, and digestion (heat treatment at 37 °C); and the cultivation of plants in soils pre-treated with diluted metal salts (vegetabilization). AMPs are manufactured in concentrated form as well as in homeopathic decimal potencies (involving successive 1:10 dilutions) and are administered as oral, rectal, vaginal, conjunctival, nasal, or percutaneous applications or by subcutaneous, intracutaneous, or intravenous injections [26]. To sum up, despite some overlap with herbal and homeopathic MPs, AMPs as a group have distinct properties that could influence their clinical safety.

Safety and effects of AMP treatment have been evaluated both on the level of the whole system and for individual MPs. The most recent comprehensive systematic review of clinical studies of AM from 2011 comprised 255 studies involving AMP treatment, with 38 studies on the level of the whole system and 217 studies of individual AMPs or small AMP groups. External validity was high but many studies had methodological shortcomings. In the reviewers' conclusion, trials of varying design and quality in a variety of diseases showed mostly good clinical outcomes, only marginal side effects, high patient satisfaction, and presumably slightly lower costs [30].

An occasion to evaluate the clinical safety of AMPs was given by the Evaluation of Anthroposophic Medicine Pharmacovigilance Network (EvaMed), a German prospective, multicenter, observational study on prescription and safety of MPs among physicians practicing AM. The EvaMed study employed a novel method for automatic extraction of anonymized core data from physicians' electronic medical records, combined with a structured approach to detection and reporting of suspected ADRs during routine outpatient care. Previous analyses from EvaMed have comprised AMP prescriptions for children during 1 year (2005 [25]), ADRs from a small MP subgroup (Asteraceae extracts [23]), and the safety of all MPs over a 5.5-year period (Jan 2004–June 2009) but without data on the subgroup of AMPs [40]. Here we present an analysis of prescription and safety of all AMPs in EvaMed. Compared with previous analyses, this report covers a longer documentation period (10 years, 2001–2010) and provides specific data on ADRs to AMPs, also in pertinent subgroups.

2 Methods

2.1 Objectives

The primary objective of this analysis was to determine the frequency of ADRs to AMPs, relative to the number of AMP prescriptions in EvaMed, for all AMPs as well as in subgroups according to age, specific AMPs or AMP groups, dosage forms, and concentrations.

Further objectives were

- to describe characteristics of the prescription of AMPs and other MPs;
- to compare the frequency of ADRs to AMPs with the frequency of ADRs to Non-AMPs (all other MPs), relative to the number of prescriptions and the number of patients prescribed AMPs and Non-AMPs, respectively.

2.2 The EvaMed Pharmacovigilance Network

EvaMed was a prospective, multicenter, observational study conducted in outpatient medical practices in Germany. Detailed descriptions of settings, participants, and data collection have been published elsewhere [24, 40]. In brief, 38 physicians (21 family physicians, 9 pediatricians, 4 internists, 4 other specialists) from 12 of 16 German Federal States participated; they had a qualification in AM, ≥ 5 years of practice experience, and used an electronic medical record system in their practice which fulfilled the technical requirements for data collection.

For each patient consultation, anonymized core data were extracted automatically from the electronic medical records: consultation date; patient age, gender, diagnoses; prescriptions of all MPs and nonmedication treatment. All physicians were obliged to link all MP prescriptions to the respective indications (diagnoses), and to document all serious ADRs as well as all ADRs of intensity III–IV to any MP. In addition, a subgroup of seven 'prescriber physicians' agreed to also document all non-serious ADRs of any intensity. Otherwise, the study was conducted under routine care conditions with ADRs identified at ordinary follow-up consultations, without any additional scheduled follow-up visits. Physicians were remunerated with 15 Euro for each ADR report but not for their regular participation; patients received no remuneration.

2.3 Eligibility Criteria for this Analysis

Eligible for this analysis were patients of physicians in the EvaMed network, with

- at least one prescription of an AMP documented in the period 1 January 2001 to 31 December 2010
- followed by at least one visit to the EvaMed physician in the period 2 January 2001 to 31 December 2011.

These criteria were selected in order to identify patients for whom ADRs to AMPs could be detected (the EvaMed documentation started in 1997, but during the years 1997–2000 only 70 MP prescriptions were documented altogether and no ADRs to AMPs occurred; therefore these years were excluded from the analysis).

2.4 Outcome Measures

The primary outcome measure for this analysis was the frequency of ADRs to AMPs [any ADR, ADRs Grade III–IV and serious ADRs (SADRs)], relative to the number of AMP prescriptions.

Secondary outcomes were:

- properties of AMPs prescribed: starting materials, manufacturing procedures, concentration of active ingredients, administration forms;
- indications for prescription of AMPs and non-AMPs;
- frequency of ADRs to non-AMPs, relative to the number of non-AMP prescriptions;
- properties of ADRs to AMPs and non-AMPs, respectively: intensity and type of symptoms; intensity, occurrence, duration, management, outcome, and seriousness of ADRs.

2.5 Documentation, Preparation, and Classification of Data

The term ‘prescription’ refers to the prescriptions of one MP, with prescriptions of more than one MP at one physician visit being counted separately.

The dataset for the present analysis was prepared by author KH by checking prescription data in the raw dataset, excluding prescriptions not referring to MPs (e.g., food supplements, body care products), prescriptions of unclassifiable MPs, and consultations without prescriptions. Subsequently, author AG checked, verified, and classified all AMP prescriptions.

2.5.1 Anthroposophic Medicinal Products (AMPs)

AMPs were defined according to the German Medicines Act [2]. For practical purposes, all MPs marketed in Germany by the manufacturers Abnoba (Pforzheim, Germany), Birken (Niefern-Öschelbronn, Germany), Helixor (Rosenfeld, Germany), Wala (Bad Boll, Germany), and Weleda (Arlesheim, Switzerland) were classified as AMPs. For AMPs, the unit of analysis was each product with a separate registration or marketing authorization, corresponding to each AMP with a separate entry in the medicines lists of the respective manufacturers. Accordingly, AMPs listed together within a separate entry but marketed in different concentrations or pack sizes were grouped together.

In EvaMed, the prescribed MPs were identified according to German National Drug Code (Pharmazentralnummer, PZN). However, the PZN codes had not been used consistently for all AMPs, with some AMPs being documented in free text by proprietary names or as abbreviations. In addition, some AMP-related PZN codes were changed during the study period. These factors brought a risk for misclassifications, in particular between anthroposophic and homeopathic or herbal MPs. In order to minimize possible classification errors, all AMP prescriptions (> 300,000) were reassessed by visual inspection of each record in the dataset, and the classification as AMPs was verified by cross-checking with a database of all AMPs

available on the market in Germany in the documentation period [the European Scientific Cooperative on Anthroposophic Medicinal Products (ESCAMP) database of AMPs, <http://www.escamp.org>, data on file].

AMPs were classified with regard to starting materials, manufacturing processes, and dosage forms, according to APC, Fourth Edition [26].

For individual AMPs mentioned in this paper, proprietary names were recoded as generic names: for AMPs manufactured from between one and three starting materials, the starting materials were listed in the order used in the medicines lists; for AMPs manufactured from four or more starting materials, the two first starting materials were listed, followed by ‘comp.’.

2.5.2 Other Medicinal Products

Non-AMPs were classified at the Research Institute Havelhöhe (German: Forschungsinstitut Havelhöhe, FIH) according to the Anatomical Therapeutic Chemical Classification System (ATC).

2.5.3 Indications

Indications for prescriptions were classified by the participating physicians according to the International Classification of Diseases, Tenth Revision (ICD-10); this was part of their routine work, since in the German Statutory Health System, physician remuneration for patient consultations requires at least one ICD-10 diagnosis for each consultation. For this analysis, diagnoses were grouped according to the ICD-10 diagnosis chapters as well as the ICD-10 diagnosis blocks.

2.5.4 Adverse Drug Reactions (ADRs)

For each ADR, the participating physicians documented

- date of onset, date of recovery;
- occurrence (once, several times, continuous);
- overall severity (classified according to the recommendations of the WHO Uppsala Monitoring Centre [8]: Grades I = mild, II = moderate, III = severe, IV = life threatening);
- symptoms;
- severity of each symptom (mild, moderate, severe);
- management (drug withdrawal, dose reduction, change of therapy, no change in drug and no additional treatment, other);
- outcome following drug withdrawal;
- rechallenge;
- outcome (recovered, not yet recovered, permanent damage, unknown, death);

- current diagnoses at onset of ADR;
- all MPs used at onset of ADR;
- date of starting MP use;
- for each MP used: physicians' assessment of causal relation to the ADR in question.

The physicians' ADR reports were assessed at FIH with classification of

- seriousness of the ADR (a serious ADR being defined according to the International Conference on Harmonization [6] as an ADR that results in death, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or is life threatening);
- causal relationship between documented ADR and the MPs used (classified according to WHO Uppsala Monitoring Centre [8]: certain, probable, possible, unlikely, conditional, unclassified, unassessable) [criteria for each category listed in Online Resource 1, see electronic supplementary material (ESM)].

The causality assessment at the FIH was performed independently by two research physicians trained in ADR evaluation, using a pre-defined case report verification form. In case of disagreement between the two research physicians, an expert team of three senior physicians and two pharmacists was consulted [40].

Symptoms of the ADRs were classified by AG according to the Medical Dictionary for Regulatory Activities (MedDRA, Version 19.0, MedDRA MSSO, McLean, VA, USA). For descriptive purposes, the frequency of ADRs was classified according to the recommendations of the Council for International Organizations of Medical Sciences (CIOMS) [5] as 'very common' ($\geq 10\%$), 'common' ($1 < x < 10\%$), 'uncommon' ($0.1 < x < 1\%$), 'rare' ($0.01 < x < 0.1\%$), and 'very rare' ($< 0.01\%$).

Each participating physician received a face-to-face training program to assist in the detection, classification, and reporting of ADRs [39].

2.6 Data Analysis

Patients fulfilling all eligibility criteria were included in the analysis. Data analysis was performed using IBM SPSS Statistics 19[®] (International Business Machines Corp., Armonk, NY, USA) and StatXact[®] 9.0.0 (Cytel Software Corporation, Cambridge, MA, USA). Missing data for individual variables were not replaced.

Analysis was descriptive. In addition, bivariate analyses of independent samples were performed using non-parametric methods: for dichotomous data, Fisher's exact test was used; for multinomial data, the Fisher–Freeman–Halton test; for rank-ordered or continuous data, Mann–

Whitney *U* test with estimation of median shift and 95% confidence interval (95% CI) according to Hodges–Lehmann. All tests were two-tailed. Significance criteria were $p < 0.05$ and 95% CI not including 0. Since this was a descriptive study, no adjustment for multiple comparisons was performed [19].

3 Results

3.1 Patients, Physicians, and Prescriptions Included in the Analysis

The dataset prepared for the analysis contained 88,682 patients with a total of 863,340 MP prescriptions. Of these, 44,662 patients with a total of 717,545 prescriptions fulfilled all eligibility criteria and were included in the analysis (Table 1).

Age groups were 0–2 years (27.1% of evaluable patients, $n = 12,065/44,662$), 3–6 years (17.9%), 7–10 years (11.1%), 11–17 years (7.5%), 18–44 years (17.7%), 45–64 years (12.6%), and ≥ 65 years (6.2%) with a median age of 8.0 years [range 0–101 years, interquartile range (IQR) 2.0–38.0 years, mean 20.1 years, standard deviation (SD) 22.9]. A total of 57.2% ($n = 25,533/44,662$) of all patients and 70.8% ($n = 11,490/16,230$) of adults (aged ≥ 18 years) were women. A total of 54.6% ($n = 24,392/44,662$) of patients were treated by family physicians, 35.4% by pediatricians, 5.6% by internists, 2.6% by dermatologists, and 1.8% by gynecologists.

Comparing patients of prescriber physicians ($n = 12,956$ patients) and of other physicians ($n = 31,617$), the patients of prescriber physicians were a median of 2.0 years younger (95% CI 2.0–2.0 years, $p < 0.001$), with a slightly lower proportion of females among all patients (55.9 vs 57.7%, $p = 0.0003$), while the gender distribution in adults did not differ significantly (72.0 vs 70.4%, $p = 0.0691$).

3.2 Documentation Period, Physician Visits, Prescriptions

For each patient, the documentation period from the first to the last physician visit during the study period was 0–11 months in 34.9% ($n = 15,598/44,662$) of patients, 12–23 months in 18.8%, 24–35 months in 14.7%, 36–47 months in 12.0%, and ≥ 48 months in 19.5%, with a median documentation period of 21.7 months (IQR 7.7–42.9 months, mean 27.1 months, SD 22.4). The number of physician visits per patient during the whole documentation period was 1–9 visits in 67.2% ($n = 30,032/44,662$) of patients, 10–19 visits in 20.6%, 20–29 visits in

Table 1 Patient inclusion into the analysis

No.	Inclusion criterion	Patients	Prescriptions			Inclusion in analysis of
			All MPs	AMPs	Non-AMPs	
1	≥ 1 MP prescription 1997–2010	88,682	863,340	336,265	527,075	
2	≥ 1 MP prescription 2001–2010	88,672	863,270	336,238	527,032	
3	No. 2 + ≥1 AMP prescription	62,956	756,757	336,238	420,519	
4	No. 3 + ≥1 physician consultation after AMP prescription (2001–2011)	44,662	717,545	311,731	405,814	MP prescriptions ADR, Grades III–IV only
5	No. 4 + patient of prescriber physician	12,956	243,147	94,734	148,413	MP prescriptions ADR, Grades I–IV

ADR adverse drug reactions, AMP anthroposophic medicinal product, MP medicinal product, non-AMP all medicinal products except AMPs

6.9%, and 30–192 visits in 5.3%, with a median of 6.0 visits (IQR 3.0–12.0 visits, mean 9.8 visits, SD 11.1) per patient.

Compared with the other patients, the patients of prescriber physicians had a longer documentation period (median difference 50.0 days, 95% CI 39.0–62.0 days, $p < 0.001$) and more physician visits (median difference 1.0 visit, 95% CI 1.0–1.0, $p < 0.001$).

During the documentation period, the 44,662 patients received a total of 717,545 MP prescriptions, of which 43.4% ($n = 311,731$) were AMP prescriptions (Table 1); the remaining prescriptions were of conventional (41.2%), homeopathic (10.0%), or herbal (5.4%) MPs (hereafter summarized as ‘non-AMPs’).

3.3 AMPs

3.3.1 Indications

Of the 311,731 AMP prescriptions, an ICD-10 diagnosis could be coded for $n = 303,725$ (97.4%) prescriptions.

Among adults aged ≥ 18 years, the most frequent ICD-10 diagnosis chapters were M00–M99 Musculoskeletal diseases (11.6%, $n = 13,404/115,057$ of prescriptions), J00–J99 Respiratory diseases (11.6%), C00–D48 Neoplasms (11.0%), and I00–I99 Cardiovascular diseases (10.3%) (Table 2). The most frequent ICD-10 Diagnosis blocks were C00–C75 Malignant neoplasms, primary, of specified sites (7.1% $n = 8143/115,046$ evaluable prescriptions), J00–J06 Acute upper respiratory infections (4.9%), M40–M54 Dorsopathies (4.6%), and I10–I15 Hypertensive diseases (4.0%).

Among children aged 0–17 years, the most frequent ICD-10 diagnosis chapters were J00–J99 Respiratory diseases (32.4% of evaluable prescriptions, $n = 61,127/188,668$), A00–B99 certain infectious and parasitic

diseases (9.9%), and R00–R99 Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (9.8%) (Table 2). The most frequent ICD-10 diagnosis blocks were J00–J06 Acute upper respiratory infections (16.5%, $n = 31,080/188,598$ evaluable prescriptions), L20–L30 Dermatitis and eczema (5.8%), H65–H75 Diseases of middle ear and mastoid (5.3%), J30–J39 Other diseases of upper respiratory tract (4.3%), and J40–J47 Chronic lower respiratory diseases (4.1%).

3.3.2 Prescriptions and Products

Of the 311,731 AMP prescriptions issued in the period 2001–2010, a total of 311,381 (99.89%) prescriptions could be matched with an AMP recorded in the ESCAMP Database of AMPs, corresponding to 1722 different AMPs, which amounts to 74.2% of a total of 2321 different AMPs marketed in Germany in this period.

Of the prescribed AMPs, 8.8% ($n = 152/1722$) were only prescribed once during the documentation period, 22.9% of AMPs had 2–9 prescriptions each, 32.2% had 10–49 prescriptions, 27.9% had 50–499 prescriptions, and 8.2% had ≥ 500 prescriptions, with a median of 25.0 prescriptions per AMP (IQR 6.0–99.3 prescriptions, mean 180.8, SD 620.0).

The numbers of AMP prescriptions per patient (repeated prescriptions of the same AMP being counted separately) were 1–4 AMP prescriptions in 54.2% ($n = 24,211/44,662$) of patients, 5–9 prescriptions in 25.3%, 10–19 prescriptions in 14.1%, and ≥ 20 prescriptions in 6.4%, with a median of 4.0 AMP prescriptions per patient (IQR 2.0–8.0 prescriptions, mean 7.0, SD 9.7). Regarding the number of AMP prescriptions per patient, the patients from prescriber physicians and other physicians did not differ significantly (median difference 0.0 prescriptions, 95% CI 0.0–0.0). A total of 57.6% ($n = 179,703/311,731$) of AMP

Table 2 Indications for prescriptions of anthroposophic medicinal products, most frequent diagnosis chapters of the International Classification of Diseases, Tenth Revision

Chapter	Code	Text	All patients		Adults (18–101 years)		Children (0–17 years)	
			<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
X	J00–J99	Diseases of the respiratory system	74,441	24.5	13,314	11.6	61,127	32.4
XVIII	R00–R99	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	24,899	8.2	6414	5.6	18,485	9.8
I	A00–B99	Certain infectious and parasitic diseases	22,529	7.4	3840	3.3	18,689	9.9
XII	L00–L99	Diseases of the skin and subcutaneous tissue	19,643	6.5	6021	5.2	13,622	7.2
V	F00–F99	Mental and behavioural disorders	18,321	6.0	9749	8.5	8572	4.5
XI	K00–K93	Diseases of the digestive system	17,947	5.9	7867	6.8	10,080	5.3
XX	V01–Y98	External causes of morbidity and mortality	17,405	5.7	2610	2.3	14,795	7.8
XIII	M00–M99	Diseases of the musculoskeletal system and connective tissue	15,569	5.1	13,404	11.6	2165	1.1
IV	E00–E90	Endocrine, nutritional and metabolic diseases	13,718	4.5	8943	7.8	4775	2.5
II	C00–D48	Neoplasms	13,393	4.4	12,703	11.0	690	0.4
IX	I00–I99	Diseases of the circulatory system	13,006	4.3	11,904	10.3	1102	0.6
VIII	H60–H95	Diseases of the ear and mastoid process	12,864	4.2	1263	1.1	11,601	6.1
XIV	N00–N99	Diseases of the genitourinary system	8687	2.9	5077	4.4	3610	1.9
XIX	S00–T98	Injury, poisoning and certain other consequences of external causes	8582	2.8	3396	3.0	5186	2.7
VII	H00–H59	Diseases of the eye and adnexa	7780	2.6	1735	1.5	6045	3.2
VI	G00–G99	Diseases of the nervous system	6236	2.1	4167	3.6	2069	1.1
Other chapters (with < 2% in each age group)			8708	2.9	2650	2.3	6055	3.2
Total evaluable diagnoses			303,725	100.0	115,057	100.0	188,668	100.0

prescriptions were issued by family physicians, 34.4% by pediatricians, 5.1% by internists, 1.8% by dermatologists, and 1.1% by gynecologists.

The 14 most frequently prescribed AMPs are listed in Table 3. Of these, eleven AMPs (all except numbers 2, 8, and 14) are usually prescribed to treat infectious diseases.

AMPs were manufactured from a single starting material (56.2%, $n = 968/1722$ AMPs), from more than one starting material (36.5%) or from one or several compositions according to the APC [26] (7.2%). Among the AMPs manufactured from a single starting material ($n = 968$), the starting material was of mineral origin in 13.8% ($n = 134/968$), of botanical origin in 46.8%, of zoological origin in 21.7%, was a chemically definable

substance in 13.0%, and a substance (metal salt or mineral) that had undergone vegetabilization in 4.6%.

Regarding concentration of active ingredients, 22.7% of evaluable AMPs ($n = 389/1711$) were manufactured in concentrated form including mother tinctures, 37.2% were in D1–D3 decimal potencies, 29.7% in D4–6, and 10.3% in potencies \geq D7. Administration forms of the prescribed AMPs were oral (44.9%, $n = 774/1,722$ AMPs), parenteral (38.4%), cutaneous (11.7%), ophthalmic (2.3%), rectal (1.7%), and other (0.9%).

3.3.3 ADRs

In the documentation period, the EvaMed physicians documented a total of 111 cases of ADRs to AMPs. Of these,

Table 3 Most frequently prescribed anthroposophic medicinal products

Order	Name of anthroposophic medicinal product	Dosage form	Prescriptions	
			N	%
1	Althaeae/Solanum comp.	Syrup	10,859	3.5
2	Belladonna/Chamomilla comp.	Suppositories	9717	3.1
3	Apis/Belladonna/Mercurius	Globuli velati	7917	2.5
4	Calendula/Eucalyptus	Nasal drops	6492	2.1
5	Apisinum/Belladonna/Cinnabar	Trituration	5412	1.7
6	Aconitum/Bryonia	Pillules	5244	1.7
7	Berberis/Prunus/Silicea	Nasal ointment	4822	1.5
8	Onopordum/Hyoscyamus/Primula	Dilution	4704	1.5
9	Levisticum	Dilution	4502	1.4
10	Bryonia/Spongia comp.	Dilution	3736	1.2
11	Petasites/Picea/Plantago	Syrup	3685	1.2
12	Pyrites	Trituration	3496	1.1
13	Pyrites/Cinnabar	Tablets	3346	1.1
14	Fragaria/Vitis	Tablets	3322	1.1
15–1722	All other products (each with < 1.0% of prescriptions)		234,127	75.2
Total number of prescriptions			311,381	100.0

Globuli velati [from Latin; globuli (little balls or spheres), velati (coated)]: coated sucrose pellets, *Pillules* small pills

the causal relation to the AMPs in question was assessed by the research scientists at FIH as certain ($n = 33$), probable ($n = 37$), possible ($n = 30$), unlikely ($n = 9$), and unclassified ($n = 2$). The 100 cases with causal relationship assessed as certain, probable or possible were classified as confirmed ADRs to AMPs and included in the ADR analysis.

The 100 confirmed ADRs were caused by a total of 109 AMPs, made up of 83 different AMPs. The cause of the ADR was one single AMP in 92.0% ($n = 92/100$) of ADRs and two or more AMPs in 8.0%. The numbers of ADRs caused by a single AMP was seven ADRs (1 AMP), five ADRs (3 AMPs), two ADRs (5 AMPs), and one ADR (77 different AMPs). In order to identify possible ADR clusters from similar AMPs, AMPs with similar starting materials were grouped together, yielding nine AMP groups. Of these, one AMP group (mistletoe products = AMPs with *Viscum album* as starting material) caused 11 ADRs and the remaining eight AMP groups caused two ADRs each.

The 100 ADRs to AMPs affected 95 patients: 90 with one ADR and five with two ADRs each. Of the 95 patients, 54 were adults aged ≥ 18 years and 51 were children aged 0–17 years, with a median age of 36.0 years (IQR 2.0–51.0 years, mean 31.9, SD 27.7). A total of 85.2% ($n = 46/54$) of adults and 42.5% ($n = 17/40$) of children with ADRs to AMPs were females.

The intensity of the ADRs to AMPs was Grade I (mild) in 50.0% ($n = 50/100$) of ADRs, Grade II (moderate) in 43.0%, Grade III (severe) in 7.0%, and Grade IV (life

threatening) in 0.0%. The number of symptoms documented for each ADR was one symptom in 74.0% ($n = 74/100$) of ADRs, two symptoms in 18.0%, three symptoms in 4.0%, four symptoms in 3.0%, and five symptoms in 1.0%, with a mean of 1.4 symptoms per ADR (SD 0.79, median 1.0, IQR 1.0–2.0). Altogether, 139 ADR symptoms were documented, with symptom intensity classified as mild in 18.7% ($n = 25/134$) of evaluable symptoms, moderate in 59.7%, and strong (German: 'stark') in 21.6%. The most common ADR symptoms, according to the MedDRA classification of System Organ Classes, were Skin and subcutaneous tissue disorders (21.6% of symptoms, $n = 30/139$), Psychiatric disorders (19.4%), Gastrointestinal disorders (17.3%), General disorders and administration site conditions (9.4%), and Respiratory, thoracic and mediastinal disorders (6.5%).

The ADRs occurred only once in 17.0% ($n = 17/100$) of ADRs, they occurred several times in 40.0%, and were continuous in 43.0%. Duration of the ADR was < 24 h in 4.3% ($n = 4/94$) of evaluable ADRs, 1–2 days in 29.8%, 3–6 days in 36.2%, 7–13 days in 10.6%, 14–30 days in 10.6%, and 31–103 days in 8.5%, with a median duration of 4.0 days (IQR 2.0–7.5 days, mean 9.8 days, SD 17.0). Management of ADRs was drug withdrawal in 80.4% ($n = 78/97$) of evaluable ADRs, dose reduction in 11.3%, and no precautions in 8.2%. Outcome of the ADR was 'recovered' in 95.0% ($n = 95/100$) of ADRs and 'not yet recovered' in 5.0%. A list of the 100 ADRs to AMPs is presented in Online Resource 2 (see ESM).

Of the 100 ADRs to AMPs, one ADR was classified by the research scientists as a SADR. Notably, for this ADR the documentation was ambiguous with respect to the seriousness. The ADR was classified as serious because of being documented as ‘life threatening at the occurrence of the symptoms’, but the overall intensity of the ADR was not documented as Grade IV (life threatening) but as Grade III (severe). Nonetheless, the classification as a SADR was upheld in the analysis. The SADR occurred in a 45-year-old woman with pre-existing ductal invasive breast cancer (ICD-10 C50.9) and pneumothorax (J93) and was caused by a mistletoe AMP (*Herba Visci albi* subsp. *abietis*, 10 mg ampoules for subcutaneous injection). Symptoms of the SADR with respective symptom severity were breathlessness (moderate), panic (strong), paraesthesia of the hands (moderate), and anxiety (strong). The causal relationship between the mistletoe AMP and the SADR was classified as probable. Concurrently with the mistletoe AMP the patient was also using an oil of *Argania spinosa*, which was judged to be unrelated to the SADR. The SADR occurred once, duration was < 24 h, management was drug withdrawal, outcome was ‘recovered’.

The frequency of ADRs to AMPs in relation to AMP prescriptions was calculated separately for ADRs of all intensities (sample restricted to patients of prescriber physicians, since only these were obliged to document ADRs of any intensity) and for ADRs with intensity Grades III–IV (all patients) (Table 4):

- ADRs of any intensity occurred in 0.071% of AMP prescriptions and in 0.502% of patients prescribed AMPs.
- ADRs with intensity Grades III–IV occurred in 0.002% of AMP prescriptions and in 0.016% of patients prescribed AMPs.

Subgroup analyses of the frequency of ADRs to AMPs in relation to AMP prescriptions were performed among patients of prescriber physicians, and comprised age (children, one group), specific AMPs or AMP groups (five groups), dosage forms (three groups), and concentrations (two groups), altogether 11 groups. With an overall ADR frequency of 0.071% of prescriptions, this frequency ranged from 0.051% (Belladonna/Chamomilla recutita, Radix comp., Suppositories) to 0.290% (Hepar sulfuris/Membrana sinuum paranasalium bovis, Ampoules) (Table 5).

3.4 Non-anthroposophic Medicinal Products (Non-AMPs)

3.4.1 Prescriptions and Products

During the documentation period, the 44,662 patients received a total of 405,814 prescriptions of non-AMPs. A

total of 11.6% ($n = 5,189/44,662$) of the patients did not have any non-AMP prescriptions, 41.5% had 1–4 non-AMP prescriptions, 20.3% had 5–9 non-AMP prescriptions, 14.8% had 10–19 non-AMP prescriptions, and 11.8% had ≥ 20 non-AMP prescriptions, with a median of 4.0 non-AMP prescriptions for each patient (IQR 1.0–10.0, mean 9.1, SD 16.3, range 0–447 non-AMP prescriptions).

The most common medication groups, according to the ATC system—first level, were Respiratory system (26.6%, $n = 107,997/405,814$ non-AMP prescriptions), alimentary tract and metabolism (14.5%), various (12.5%), nervous system (9.8%), cardiovascular system (8.2%), dermatologicals (6.4%), and Anti-infectives for systemic use (6.2%). The most common second-level groups were R05 Cough and cold preparations (13.7%, $n = 55,475/405,814$ prescriptions), V60 Homeopathic MPs¹ (9.5%), R01 Nasal preparations (6.8%), J01 Antibacterials for systemic use (5.2%), N02 analgesics (4.8%), R03 Drugs for obstructive airway diseases (4.2%), and M01 anti-inflammatory and antirheumatic products (3.2%).

3.4.2 Adverse Drug Reactions to Non-AMPs

A total of 695 cases of ADRs to non-AMPs were documented, of which 682 ADRs were confirmed by the FIH research scientists. The 682 ADRs affected 637 patients and were caused by 775 different non-AMPs. Of the 682 ADRs, 29 ADRs (4.3%) were SADRs, affecting 29 different patients.

The most common causes of ADRs to non-AMPs, according to the ATC system—second level, were J07 Vaccines (46.0%, $n = 353/768$ evaluable non-AMPs), J01 Antibacterials for systemic use (24.1%), and R05 cough and cold preparations (7.4%).

3.5 Comparison of AMPs and Non-AMPs

3.5.1 Indications

Comparing the most common indications for AMPs versus non-AMPs in adults (Online Resource 3, see ESM) and in children (Online Resource 4, see ESM), respectively, the relative frequency of ICD-10 diagnosis chapters were similar: the only differences above 3% (absolute percentages) were for I00–I99 Diseases of the circulatory system (10.3% of AMP prescriptions vs 15.8% of non-AMP prescriptions), and E00–E90 Endocrine, nutritional and metabolic diseases (7.8% vs 11.5%) in adults and A00–B99

¹ V60 is used as the ATC code in Germany for homeopathic and anthroposophic MPs, but in this subgroup analysis of non-AMPs only the homeopathic MPs were included.

Table 4 Frequency of adverse drug reactions in relation to prescriptions

	All age groups, 0–101 years		Children, 0–17 years	
	All AMPS	All non-AMPS	All AMPS	All non-AMPS
Patients of prescriber physicians				
ADRs, Grades I–IV	67	612	37	574
Prescriptions	94,734	148,413	59,265	120,638
ADRs/prescriptions (%)	0.071	0.412	0.062	0.453
Patients with ADR	65	569	36	532
Patients with prescriptions	12,956	11,652	9,162	8,550
Patients with ADR/with prescriptions (%)	0.502	4.88	0.393	6.22
Patients of all physicians				
ADRs, Grades III–IV only	7	82	2	59
Prescriptions	311,731	405,814	193,851	243,560
ADRs/prescriptions (%)	0.002	0.020	0.001	0.024
Patients with ADR	7	82	2	59
Patients with prescriptions	44,662	39,473	28,355	25,587
Patients with ADR/with prescriptions (%)	0.016	0.208	0.007	0.231

ADRs adverse drug reactions, AMPs anthroposophic medicinal products, non-AMPs all medicinal products except AMPs

Table 5 Frequency of adverse drug reactions to anthroposophic medicinal products, subgroup analyses

All age groups, patients of prescriber physicians	ADRs, all grades	Prescriptions	ADRs/prescriptions (%)	Patients with ADR	Patients with prescription	Patients with ADR/with prescription (%)
All patients of prescriber physicians	67	94,734	0.071	65	12,956	0.502
Specific AMPs or AMP groups*						
Mistletoe AMPs, ampoules	3	4687	0.064	3	551	0.544
Fragaria/Vitis, tablets	7	3322	0.211	7	1771	0.395
Apis/Belladonna, dilution	5	2572	0.194	4	2116	0.189
Belladonna/Chamomilla recutita, Radix comp., suppositories	5	9717	0.051	5	5100	0.098
Hepar sulfuris/Membrana sinuum paranasalium bovis, ampoules	5	1722	0.290	4	1069	0.374
Dosage forms						
Injectables (parenteral)	27	16,378	0.165	24	3633	0.661
Oral	67	53,444	0.125	58	11,660	0.497
Local (cutaneous, ophthalmic, rectal, other)	15	24,875	0.060	15	8679	0.173
Concentrations						
Non-diluted, D1–D3	84	76,164	0.110	78	12,635	0.617
≥ D4	24	17,249	0.139	23	6436	0.357

ADRs adverse drug reactions, AMPs anthroposophic medicinal products

*AMP groups: Mistletoe AMPs

Certain infectious and parasitic diseases (9.9 vs 13.1%) in children.

3.5.2 Features of ADRs

Comparing the relative frequencies (not the absolute numbers) of features of ADRs to AMPs and non-AMPs,

respectively, the ADRs to AMPs had significantly lower severity (Grade I severity being a feature of 50.0% for AMPs vs 20.4% for non-AMPs, $p < 0.0001$), were less likely to occur continuously (43.0 vs 79.4%, $p < 0.0001$), and were more often treated with drug withdrawal (80.4 vs 56.5%, $p < 0.0001$). Regarding seriousness, duration, and outcome of ADRs, the relative frequencies did not show

any significant differences between the two groups (Online Resource 5, see ESM).

3.5.3 Frequencies of ADRs, Relative to Prescription

Among patients of prescriber physicians, the frequency of ADRs of any severity, relative to prescription rate, was significantly lower for AMPs (0.071%, $n = 67$ ADRs/94,734 prescriptions) than for non-AMPs (0.412%, $n = 612$ ADRs/148,413 prescriptions) [odds ratio (OR) for ADRs to AMPs vs ADRs to non-AMPs 0.17, 95% CI 0.13–0.22, $p < 0.0001$] (Table 4). The frequency of ADRs to AMPs and non-AMPs, respectively, was analyzed in the three largest diagnostic subgroups in adults and children, respectively. In all six subgroups, ADRs to non-AMPs were more frequent than ADRs to AMPs; differences were significant for two subgroups in children: ICD Chapters A00–B99 (infectious diseases, OR for ADRs to AMPs vs ADRs to non-AMPs 0.11, 95% CI 0.02–0.33, $p < 0.0001$) and J00–J99 (respiratory diseases, OR 0.34, 95% CI 0.21–0.53, $p < 0.0001$) (Table 6).

Among all patients, the corresponding frequency of ADRs of Grades III–IV was 0.002% ($n = 7$ ADRs/311,731 prescriptions) and 0.020% (82 ADRs/405,814 prescriptions) for AMPs and non-AMPs, respectively (OR 0.11, 95% CI 0.05–0.24, $p < 0.0001$) (cf. also Table 6 for diagnostic subgroups). The frequency of SADR was 0.0003% (1 SADR/311,731 prescriptions) for AMPs and 0.0071% (29 SADR/405,814 prescriptions) for non-AMPs (OR 0.045, 95% CI 0.001–0.271, $p < 0.0001$).

4 Discussion

4.1 Main Findings

This is the largest analysis of AMP safety in a prospective patient cohort so far: 44,662 patients from the EvaMed pharmacovigilance network with a total of 311,731 AMP prescriptions were followed up for a mean of 27 months. ADRs to AMPs were rare (0.071% of AMP prescriptions), high intensity ADRs and SADR were very rare (0.002% and 0.0003%, respectively). Among 11 analyzed subgroups, the highest ADR frequency was still in the ‘uncommon’ range (0.290% of prescriptions for one specific AMP). Compared with ADRs from non-AMPs, ADRs to AMPs were less frequent (OR for ADRs to AMPs and non-AMPs, respectively, ranging from 0.045 for SADR to 0.17 for any ADR).

4.2 Strength and Limitations

An important strength of the EvaMed study and this analysis lies in the automatic data extraction from the electronic medical records of the participating physicians.

This procedure enabled the inclusion of all MP prescriptions to all patients without any selection, and helped to achieve a very large sample of children and adults with MP prescriptions, increasing the power to detect rare ADRs as well as ADRs in subgroups of interest. Further strengths include the representation of outpatient care in 12 of 16 German Federal States; a long documentation period; the coverage of three-quarters of all AMPs marketed in Germany in the study period; and a comprehensive and stringent assessment of ADRs to all MPs.

Notably, the study was not designed to test MP effectiveness, hence no clinical outcomes were documented. The safety documentation was restricted to suspected ADRs to any MP; an ongoing documentation and assessment of all adverse events following drug prescription in this very large patient sample was not feasible. Therefore, unusual or unexpected ADRs to AMPs or other MPs, not recognized as such, cannot be excluded (this question has been assessed in other studies of AMP therapy, e.g. [22, 41]). Also, the study documentation did not include any scheduled follow-up visits; accordingly, detection of ADRs was dependent on patients returning for subsequent consultations or otherwise notifying the physicians. Hence, underreporting of suspected ADRs by the patients to their physicians cannot be excluded in this study, similar to real-world medical practice.

Considering the long documentation period of 27 months on average per physician and the inclusion of all patients seen by the physicians in the study project, ADR underreporting by the physicians is also a potential limitation. Precautions against physician underreporting were taken at the levels of study organization as well as analysis: Each physician was individually trained in ADR reporting [39]. Furthermore, the obligation of physicians to report ADRs was graded: All physicians reported the kind of ADRs that are easily noticed and where awareness may be highest, that is, serious ADRs and ADRs of high intensity. The additional documentation of nonserious ADRs of mild to moderate intensity was limited to a smaller group of ‘prescriber’ physicians who had explicitly agreed to do so. In order to have an unbiased estimate of the frequency of these ADRs in relation to prescription rates, the sample for that analysis was restricted to patients of the prescriber physicians.

For causality assessment of adverse events and suspected ADRs, different criteria and instruments are available: in a systematic review from 2008 a total of 34 different methods were identified. Out of 13 possible criteria for assigning causality (12 listed + ‘other’), most methods used five or six. The authors concluded that so far, no method had shown consistent and reproducible measurement of causality; therefore, no single method was universally accepted [1]. In EvaMed, the causality assessment system of the Uppsala Monitoring Centre of the

Table 6 Frequency of adverse drug reactions in relation to prescriptions in the three largest diagnostic subgroups in adults and children, respectively

	Adults (≥ 18 years)						Children (0–17 years)					
	Circulatory (I00–I99) ^a		Musculoskeletal (M00–M99) ^a		Respiratory (ICD-10 J00–J99) ^a		Respiratory (ICD-10 J00–J99) ^a		Infectious (A00–B99) ^a		Symptoms, signs, findings (R00–R99) ^a	
	AMPs	Non-AMPs	AMPs	Non-AMPs	AMPs	Non-AMPs	AMPs	Non-AMPs	AMPs	Non-AMPs	AMPs	Non-AMPs
Patients of prescriber physicians												
ADRs, Grades I–IV	8	18	11	10	9	13	24	154	3	82	5	22
Prescriptions	3614	4665	4145	2995	3922	2784	18,261	40,217	6690	19,926	6162	10,759
ADRs/prescriptions (%)	0.221	0.386	0.265	0.334	0.229	0.467	0.131	0.383	0.045	0.412	0.081	0.204
Patients with ADR	7	17	11	10	8	12	23	145	3	80	5	22
Patients with prescriptions	878	678	1013	716	1670	1159	6429	6641	3732	5711	3463	4384
Patients with ADR/with prescriptions (%)	0.80	2.51	1.09	1.40	0.48	1.04	0.36	2.18	0.08	1.40	0.14	0.50
Patients of all physicians												
ADRs, Grades III–IV only	2	7	1	7	2	7	2	14	0	6	0	2
Prescriptions	11,904	25,408	13,404	17,250	13,314	14,681	61,127	74,714	18,689	31,078	18,485	24,343
ADRs/prescriptions (%)	0.017	0.028	0.007	0.041	0.015	0.048	0.003	0.019	0.000	0.019	0.000	0.008
Patients with ADR	2	7	1	7	2	7	2	14	0	6	0	2
Patients with prescriptions	3613	3490	4004	3662	5795	4626	19,414	17,187	10,082	11,928	9906	10,334
Patients with ADR/with prescriptions (%)	0.06	0.20	0.02	0.19	0.03	0.15	0.01	0.08	0.000	0.05	0.000	0.02

ADRs adverse drug reactions, AMPs anthroposophic medicinal products, non-AMPs all other medicinal products

^aInternational Classification of Diseases (ICD) diagnosis chapter

World Health Organization (WHO-UMC) was used [8]. For each of the categories in the WHO-UMC system, criteria are listed (Online Resource 1, see ESM) but, unlike for many other instruments [1], not formalized into a fixed algorithm or probability score. The WHO-UMC method has been criticized for being subjective and imprecise [1], since it is to some extent based on expert opinion. However, other presumed ‘objective’ instruments may require subjective judgment (as noted for the question on “previous conclusive reports on this reaction?” [31] in the Narajo score [34]). Algorithms lack flexibility [1] and may lead to loss of relevant information [33], while probability scores presuppose that the certainty of causality assessment can be reliably modeled in a mathematical formula, which is difficult to prove. Comparisons of different instruments are complicated by their performance varying in different situations (e.g. design [18], setting [33], types of ADR [31], professional background of the raters [18]). In EvaMed, causality assessment was performed independently by two different research scientists, with disagreements solved by discussion in an interdisciplinary expert team, which would expect to reduce any bias from subjectivity. An advantage of the WHO-UMC method categories is their widespread

use amongst others in the WHO Drug Monitoring Programme, with currently 127 participating countries worldwide [10].

The EvaMed physicians were paid 15 Euro per filed ADR report as compensation for the time spent investigating the suspected ADR, interviewing the patient, and filing the report [40]. These physicians worked within the German Statutory Health System, where they are paid for each patient-related service provided, while the physicians pay all expenses for office, equipment, salaries to assistants, etc. The sum of 15 Euro was calculated to be high enough for EvaMed physicians not to suffer a financial loss from ADR reporting, but low enough for this activity not to be financially attractive. Thus, neither over-reporting bias due to economic gain nor under-reporting bias due to avoidance of economic strain would be expected from this provision. Compared with total study costs, the costs for reimbursing physicians for ADR were negligible; such a fee could be considered also for other pharmacovigilance projects.

Some items of interest for safety assessment were not documented, such as the number of days or applications for each prescribed MP, and patient or caregiver compliance

with prescriptions. Furthermore, the documentation of demographics and baseline morbidity was limited to age, gender, and diagnoses, offering limited possibilities to analyze factors associated with the occurrence of ADRs, apart from MP-related variables.

4.3 Comparison with Other Studies, Interpretation

The findings from this analysis of AMP safety can be compared with two other AMP safety analyses from prospective cohorts [21, 22], for which patient-level data are available. All three studies were from outpatient, mainly primary care, medical practices in Western countries and all included comprehensive ADR documentation with causality assessment by independent research staff. For the two comparisons, the samples of EvaMed and the comparison studies were restricted in order to have identical age and diagnosis groups (details in Online Resources 6 and 7, see ESM):

- The International Integrative Primary Care Outcomes Study (IIPCOS) concerned patients with acute respiratory and ear infections; therefore, our comparison was restricted to EvaMed patients with the same diagnoses. ADRs to AMPs were similarly frequent in IIPCOS and EvaMed (0.28 vs 0.34% of AMP users). Notably, the absolute number of ADRs to AMPs in IIPCOS was low ($n = 2$); thus, the EvaMed analysis confirms the results from IIPCOS for this patient group (Online Resource 6).
- The Anthroposophic Medicine Outcomes Study (AMOS) had a much older patient sample than EvaMed and 80% of all confirmed ADRs to AMPs in AMOS occurred in adults; therefore our comparison was restricted to adult patients. In AMOS, patients and physicians repeatedly filled in follow-up questionnaires (EvaMed had no scheduled follow-up documentation) and suspected ADRs were documented independently by patients as well as physicians (EvaMed: physicians only), with one-third of the confirmed ADRs to AMPs in the AMOS analysis being documented by patients only [22]. ADRs to AMPs were approximately 4.5 times more frequent in AMOS than in EvaMed (3.08% vs 0.69% of AMP users); this difference could be explained by the above-mentioned differences in study documentation (Online Resource 7).

SADRs to AMPs were very rare in EvaMed (0.0022% of 44,662 users) and did not occur in the two other studies.

Also, the frequency of ADRs to non-AMPs in this study can be compared with other studies. Focusing on children aged 0–17 years, who made up almost two-thirds (63.6%) of patients in this analysis, and matching for study design (observational studies), setting (outpatients), and

denominator (number of patients exposed to a drug), we identified four studies from a recent, very comprehensive systematic review of ADRs in children [37]. The frequency of ADRs to non-AMPs among drug-exposed pediatric outpatients was 6.22% in EvaMed (Table 4) and 5.92% in the four studies [17, 28, 35, 36] (weighted mean: $n = 404/6824$ patients, cf. Fig. 3 in [37]). Withstanding the limitations of such comparisons from confounding by other variables (e.g., indications and type of MPs), the incidence of ADRs to non-AMPs in EvaMed seems to be in the same order of magnitude as in other observational studies.

The primary focus of this analysis was safety of AMPs, which is from a regulatory perspective a distinct topic. However, with AM being an integrative system of medicine, AM physicians prescribe not only AMPs in their patient care but also other MPs. Accordingly, a broader, public health focus on the safety of all MPs prescribed is also relevant. In this analysis, ADRs to any MP were uncommon (0.28% of prescriptions, $n = 679/243,147$). Notably, within this broad focus, ADRs to AMPs were less frequent than ADRs to non-AMPs, with OR for ADRs to AMPs and non-AMPs, respectively, of 0.17. This difference was even larger for ADRs of high intensity (OR 0.11) and SADRs (OR 0.045). Notably, the two types of MPs had relatively similar indications. Therefore, in routine outpatient care, risks of AMP use seem to be low, not only in absolute terms but also in comparison to risks from other types of MPs used by the patients.

5 Conclusion

This analysis of more than 40,000 outpatients in a prospective cohort study confirms previous findings from other studies with < 1000 patients [21, 22]: ADRs to AMPs were rare (0.071% of AMP prescriptions), ADRs of high intensity and SADRs were very rare (0.002 and 0.0003%, respectively).

Acknowledgements The EvaMed study was funded by the Software AG-Stiftung. This analysis and publication was commissioned by the European Scientific Cooperative on Anthroposophic Medicinal Products (<http://www.escamp.org>) with financial support from foundations (Christophorus-Stiftung, Damus-Donata-Stiftung, Ekhagastiftelsen, Mahle-Stiftung, Software AG-Stiftung) and manufacturers (Wala, Weleda). The sponsors had no influence on the planning or conduct of the EvaMed study; the collection, preparation, analysis or interpretation of data for this paper; nor on the preparation, review or approval of the manuscript.

Author Contributions HJH, AG and KH wrote the analysis plan. AG and KH prepared data. AG analyzed data. HM designed and was director of the EvaMed study. HJH was principal author, had access to all data, and is guarantor. All authors approved the final manuscript.

Compliance with Ethical Standards

Quality assurance, adherence to guidelines, approval The EvaMed study was based on secondary data provided by physicians and adhered to the recommendations for good practice in secondary data analysis from the German Working Group on the Collection and Use of Secondary Data [38]. The study was approved by the responsible data security official. As all data were collected anonymously without any documentation on the part of the patients, and since the study did not entail any intervention or experimental research, no ethical approval was required [16]. This report followed the STROBE guidelines for reporting of observational studies [42].

Competing interests In the past 3 years, HJH has received research grants from manufacturers of AMPs (Wala, Weleda). Since 2012 (after the EvaMed study was completed), HM has been member of the Board of Directors of Weleda, manufacturer of AMPs. AG and KH declare that they have no competing interests.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Agbabiaka TB, Savovic J, Ernst E. Methods for causality assessment of adverse drug reactions: a systematic review. *Drug Saf.* 2008;31:21–37.
- Anonymous (2014) Arzneimittelgesetz in der Fassung der Bekanntmachung vom 12. Dezember 2005 (BGBl. I S. 3394), das zuletzt durch Artikel 2a des Gesetzes vom 27. März 2014 (BGBl. I S. 261) geändert worden ist. Bundesministerium der Justiz und für Verbraucherschutz, Bonn.
- Anonymous (2016) EMA pharmacovigilance system manual. Version 1.2. In: European Medicines Agency, p. 26.
- Anonymous. Facts and figures on Anthroposophic Medicine worldwide. Brussels: IVAA-International Federation of Anthroposophic Medical Associations; 2012.
- Anonymous (2007) Frequency of adverse drug reactions (CIOMS). In: Definitions. World Health Organization, Geneva, p. 10.
- Anonymous (1997) ICH harmonised tripartite guideline. Guideline for good clinical practice. In: International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use (ICH), Geneva.
- Anonymous. Safety issues in the preparation of homeopathic medicines. Geneva: World Health Organization; 2009.
- Anonymous. Safety monitoring of medicinal products: guidelines for setting up and running a pharmacovigilance centre. Uppsala: Uppsala Monitoring Centre; 2000.
- Anonymous. The safety of medicines in public health programmes: pharmacovigilance an essential tool. Geneva: World Health Organization; 2006.
- Anonymous. The story of UMC and the WHO Programme. Uppsala: Uppsala Monitoring Centre-WHO Collaborating Centre for International Drug Monitoring; 2017.
- Anonymous. WHO guidelines on safety monitoring of herbal medicines in pharmacovigilance systems. Geneva: World Health Organization; 2004.
- Baars EW, Hamre HJ. Whole medical systems versus the system of conventional biomedicine: a critical review of similarities, differences and factors that promote the integration process. *Evid Based Complement Altern Med eCAM.* 2017. doi:10.1155/2017/4904930.
- Baghel M. The national pharmacovigilance program for ayurveda, siddha and unani drugs: current status. *Int J Ayurveda Res.* 2010;1:197–8.
- Bornhöft G, Matthiessen PF. Homeopathy in healthcare: effectiveness, appropriateness, safety, costs. Berlin: Springer; 2011.
- Buchholzer ML, Werner C, Knoess W. Current concepts on integrative safety assessment of active substances of botanical, mineral or chemical origin in homeopathic medicinal products within the European regulatory framework. *Regul Toxicol Pharmacol.* 2013;68:193–200.
- Chenot JF, Heidenreich R. The role of ethics committees in primary care research projects. *Z Allg Med.* 2004;80:279–81.
- Cirko-Begovic A, Vrhovac B, Bakran I. Intensive monitoring of adverse drug reactions in infants and preschool children. *Eur J Clin Pharmacol.* 1989;36:63–5.
- Davies EC, Rowe PH, James S, et al. An investigation of disagreement in causality assessment of adverse drug reactions. *Pharm Med.* 2011;25:17–24.
- Feise RJ. Do multiple outcome measures require p-value adjustment? *BMC Med Res Methodol.* 2002;2:8.
- Fonnebø V, Grimsgaard S, Walach H, et al. Researching complementary and alternative treatments—the gatekeepers are not at home. *BMC Med Res Methodol.* 2007;7:7.
- Hamre HJ, Glockmann A, Fischer M, et al. Use and safety of anthroposophic medications for acute respiratory and ear infections: a prospective cohort study. *Drug Target Insights.* 2007;2:209–19.
- Hamre HJ, Witt CM, Glockmann A, et al. Use and safety of anthroposophic medications in chronic disease: a 2-year prospective analysis. *Drug Saf.* 2006;29:1173–89.
- Jeschke E, Ostermann T, Luke C, et al. Remedies containing Asteraceae extracts: a prospective observational study of prescribing patterns and adverse drug reactions in German primary care. *Drug Saf.* 2009;32:691–706.
- Jeschke E, Ostermann T, Tabali M, et al. Diagnostic profiles and prescribing patterns in everyday anthroposophic medical practice—a prospective multi-centre study. *Forsch Komplementmed.* 2009;16:325–33.
- Jeschke E, Ostermann T, Tabali M, et al. Anthroposophic medicine in pediatric primary care—a prospective, multicentre observational study on prescribing patterns. *Altern Ther Health Med.* 2011;17:18–28.
- Jones G, Maier J, Pedersen P, et al., editors. Anthroposophic Pharmaceutical Codex APC. 4th ed. Dornach: International Association of Anthroposophic Pharmacists; 2017.
- Jong MC, Jong MU, Baars EW. Adverse drug reactions to anthroposophic and homeopathic solutions for injection: a systematic evaluation of German pharmacovigilance databases. *Pharmacoepidemiol Drug Saf.* 2012;21:1295–301.
- Kaushal R, Goldmann DA, Keohane CA, et al. Adverse drug events in pediatric outpatients. *Ambul Pediatr Off J Ambul Pediatr Assoc.* 2007;7:383–9.
- Kienle GS, Albonico HU, Fischer L, et al. Complementary therapy systems and their integrative evaluation. *Explore (NY).* 2011;7:175–87.
- Kienle GS, Glockmann A, Grugel R, et al. Klinische Forschung zur Anthroposophischen Medizin—Update eines Health Technology Assessment-Berichts und Status Quo [Clinical research on anthroposophic medicine: update of a health technology assessment report and status quo]. *Forsch Komplementmed.* 2011;18:269–82.

31. Mittal N, Gupta MC. Comparison of agreement and rational uses of the WHO and Naranjo adverse event causality assessment tools. *J Pharmacol Pharmacother*. 2015;6:91–3.
32. Mol C, Roemer F, Scheer R, et al. Arzneimittel [überarbeitete Monographien der Kommission C]. In: Glöckler M, editor. *Anthroposophische Arzneitherapie für Ärzte und Apotheker*. Stuttgart: Wissenschaftliche Verlagsgesellschaft; 2008.
33. Mouton JP, Mehta U, Rossiter DP, et al. Interrater agreement of two adverse drug reaction causality assessment methods: a randomised comparison of the Liverpool Adverse Drug Reaction Causality Assessment Tool and the World Health Organization-Uppsala Monitoring Centre system. *PLoS One*. 2017;12:e0172830.
34. Naranjo CA, Busto U, Sellers EM, et al. A method for estimating the probability of adverse drug reactions. *Clin Pharmacol Ther*. 1981;30:239–45.
35. Rebelo Gomes E, Fonseca J, Araujo L, et al. Drug allergy claims in children: from self-reporting to confirmed diagnosis. *Clin Exp Allergy J Br Soc Allergy Clin Immunol*. 2008;38:191–8.
36. Sanz E, Boada J. Adverse drug reactions in paediatric outpatients. *Int J Clin Pharmacol Res*. 1987;7:169–72.
37. Smyth RM, Gargon E, Kirkham J, et al. Adverse drug reactions in children—a systematic review. *PLoS One*. 2012;7:e24061.
38. Swart E, Gothe H, Geyer S, et al. Good practice of secondary data analysis (GPS): guidelines and recommendations. *Gesundheitswesen*. 2015;77:120–6.
39. Tabali M, Jeschke E, Bockelbrink A, et al. Educational intervention to improve physician reporting of adverse drug reactions (ADRs) in a primary care setting in complementary and alternative medicine. *BMC Public Health*. 2009;9:274.
40. Tabali M, Ostermann T, Jeschke E, et al. Adverse drug reactions for CAM and conventional drugs detected in a network of physicians certified to prescribe CAM drugs. *J Manag Care Pharm*. 2012;18:427–38.
41. Tröger W, Galun D, Reif M, et al. *Viscum album* [L.] extract therapy in patients with locally advanced or metastatic pancreatic cancer: a randomised clinical trial on overall survival. *Eur J Cancer*. 2013;49:3788–97.
42. Von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med*. 2007;147:573–7.
43. Zhang L, Yan J, Liu X, et al. Pharmacovigilance practice and risk control of traditional Chinese medicine drugs in China: current status and future perspective. *J Ethnopharmacol*. 2012;140:519–25.