



# Searching for the Genus Epidemicus in Chinese Patients: Findings from the Clifical COVID-19 Clinical Case Registry

Alexander Tournier<sup>1,2</sup> Yvonne Fok<sup>3</sup> Robbert van Haselen<sup>4</sup> Aaron To<sup>3</sup>

<sup>1</sup>Institute of Complementary and Integrative Medicine, Faculty of Medicine, University of Bern, Bern, Switzerland

<sup>2</sup>Homeopathy Research Institute, London, England

<sup>3</sup>Hong Kong Association of Homeopathy, People's Republic of China

<sup>4</sup>International Institute for Integrated Medicine, Kingston, United Kingdom

Address for correspondence Robbert van Haselen, MSc, International Institute for Integrated Medicine, Suite 467, 29 Castle Street, Kingston KT1 1DN, United Kingdom (e-mail: vanhaselen@intmedi.com).

Homeopathy 2023;112:30–39.

## Abstract

**Background** The Clifical COVID-19 Support Project is an innovative international data collection project aimed at tackling some of the core questions in homeopathy. This paper reports on the further investigation of the genus epidemicus concept during the first wave of the pandemic in the Chinese population.

**Methods** The design is an observational clinical case registry study of Chinese patients with confirmed or suspected coronavirus disease 2019 (COVID-19). The symptoms were prospectively collected via a 150-item questionnaire. The concept of genus epidemicus, including the role of treatment individualization, was investigated by analyzing whether presenting symptoms clustered into distinct groups. Two standard statistical analysis techniques were utilized: principal component analysis for extracting the most meaningful symptoms of the dataset; the *k*-means clustering algorithm for automatically assigning groups based on similarity between presenting symptoms.

**Results** 20 Chinese practitioners collected 359 cases in the first half of 2020 (766 consultations, 363 prescriptions). The cluster analysis found two to be the optimum number of clusters. These two symptomatic clusters had a high overlap with the two most commonly prescribed remedies in these sub-populations: in cluster 1 there were 297 prescriptions, 95.6% of which were *Gelsemium sempervirens*; in cluster 2 there were 61 prescriptions, 95.1% of which were *Bryonia alba*.

**Conclusion** This is the first study to investigate the notion of genus epidemicus by using modern statistical techniques. These analyses identified at least two distinct symptom pictures. The notion of a single COVID-19 genus epidemicus did not apply in the studied population.

## Keywords

- ▶ genus epidemicus
- ▶ COVID-19
- ▶ clinical case registry

## Introduction

The first coronavirus disease 2019 (COVID-19) cases in China were reported in December 2019. The first wave of acute

COVID-19 cases occurred in China and lasted from December 2019 to March 2020. As of March 18, 2020, 13,415 confirmed cases and 120 deaths were reported outside Hubei province, the epicenter of the outbreak.<sup>1</sup> Hong Kong

received

December 8, 2021

accepted after revision

March 7, 2022

article published online

October 1, 2022

DOI <https://doi.org/10.1055/s-0042-1750380>.

ISSN 1475-4916.

© 2022. The Faculty of Homeopathy. All rights reserved.

This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

is geographically located in the province of Guangdong which, after Hubei, has been the region of China second-most affected by the first wave of the COVID-19 pandemic. It was managed mainly via strict lockdowns and quarantines of affected cases. In the absence of any established effective treatment, many patients sought adjunctive homeopathic treatment. For instance, it is estimated that 2 to 3% of the Hong Kong population used homeopathy as an adjunctive treatment alongside conventional and/or traditional Chinese medicines (*Living Homeopathy Ltd, internal data*).

Living Homeopathy is a homeopathic clinic, based in Hong Kong, offering treatment to a large number of patients in mainland China as well as to the Hong Kong and Macau Special Administrative Region populations. Approximately 100 practitioners provide homeopathic treatment at this clinic. Initial experiences, focusing on the homeopathic clinical features of 18 Chinese patients, were reported in 2020.<sup>2</sup>

Homeopathy has a long history of use in epidemic disease<sup>3,4,5,6,7</sup> and the COVID-19 pandemic has led to a sharp increase in demand for homeopathy as an adjunctive treatment. This is also borne out by multiple publications of surveys,<sup>8,9,10</sup> case reports,<sup>11</sup> case series,<sup>2,12,13</sup> prospective and retrospective cohort studies,<sup>14,15,16</sup> as well as research projects and initiatives.<sup>17,18,19,20</sup>

The foundations for the Clinical File Collection (Clificol) case registry were laid in the 1990s, and in the first stage of the development of the Clificol platform, use was restricted to users of the RadarOpus software platform. The second stage in the development of the Clificol platform took place in 2020 when access was broadened to all registered homeopathic practitioners. This was achieved by additionally making the Clificol platform available online (<https://www.clificol.net>). The platform is supported by all major homeopathic organizations worldwide, representing 197 associations in total. It is free-to-access for homeopathic practitioners, who are encouraged to share their cases and experiences.

The urgency and severity of the COVID-19 pandemic triggered the worldwide launch of the Clificol COVID-19 Support Project on 11<sup>th</sup> May 2020. At the time of submission of this paper there were 1,189 cases in the database. A significant portion of the initially collected data came from the first COVID-19 wave in China. Initial experiences with the Clificol platform were published<sup>21,22</sup> and this included experiences in China. The Chinese data have the additional peculiarity that most of the data were collected via administering a questionnaire. Data added subsequently from other parts of the world to the Clificol platform were mostly *not* based on the use of questionnaires. We therefore decided to report on the “first wave” Chinese data as a more homogeneous patient population in terms of nationality as well as the data collection method.

The “genus epidemicus” is the remedy found to be most effective for a particular epidemic once data have been gathered from several cases.<sup>23</sup> This concept was mentioned by Hahnemann in paragraph 241 of his “Organon”, and based on his experiences with the use of *Belladonna* in a scarlet

fever epidemic in 1799.<sup>24</sup> From the perspective of treatment individualization, it is the homeopathic remedy that best matches the totality of the symptom picture of the epidemic disease concerned. A common assumption is that each epidemic is likely to have a single genus epidemicus remedy.

In this paper we report on the nature and homeopathic treatment of COVID-19 in the first wave in the Chinese population. In particular, we aimed to investigate the possible presence and nature of a COVID-19 genus epidemicus. This article will focus on the genus epidemicus question in the Chinese population: subsequent articles will explore the genus epidemicus question in other parts of the world as well as the role of treatment individualization.

## Materials and Methods

A retrospective analysis of prospectively collected questionnaire based COVID-19 clinical case data took place. The recruitment and treatment of patients was organized by the Living Homeopathy Clinic in Hong Kong. A team of 20 practitioners was involved in the co-ordination of the recruitment, questionnaire administration and treatment of patients. A small minority of patients were recruited while attending the Clinic. All the other patients were recruited and treated outside the clinic, using videoconferencing or other appropriate means of communication. Acute COVID-19 cases from China, treated during the first wave of COVID-19 between 1<sup>st</sup> February 2020 and 30<sup>th</sup> December 2020, were eligible. Eligible patients needed to have at least one of the following diagnostic criteria as described in the 7th edition of the diagnosis and treatment protocol in China<sup>25</sup>: (1) reverse transcription polymerase chain reaction (RT-PCR) positive for 2019-novel coronavirus (2019-nCoV); (2) serology test positive for IgM/IgG (immunoglobulin M/immunoglobulin G) specific for COVID-19; (3) suspected COVID-19. The latter was defined as follows: fever, chills, or respiratory symptoms related to an outbreak cluster, where a cluster was three or more symptomatic patients, identified in a 2-week period in a small unit such as family, office or school class, in which one or more cases were RT-PCR positive.

A questionnaire for collecting COVID-19 related symptoms was developed by the Hong Kong team. It was based on a pre-existing questionnaire utilized in previous flu seasons, and then modified in line with the information available on COVID-19 at the time. It included the mandatory assessment of several clinical symptoms such as fever, chills, cough, headache, intestinal symptoms, weakness, and some further homeopathic symptoms and modalities such as thirst, perspiration and weather influences. For some symptoms, when present, additional symptoms were assessed. For instance, if patients reported cough, there were additional questions on how the cough started, the type of cough and questions on some cough-related modalities. Only patients who provided informed consent to completing the questionnaire were eligible for inclusion in the data analysis. All data were anonymized during uploading to the online platform. An overview of the symptom assessment is given in ►Table 1.

**Table 1** Overview of the symptom questionnaire that was completed by the patients

Symptom type	Mandatory symptoms assessed	Additional symptoms assessed	Comment
<i>Main/presenting symptoms</i>	Fever		
	Chills	Characteristic of chills Modalities	
	Cough	How did cough start Associated symptoms Modalities Type of cough Nature of sputum	
	Lethargy		
	Shortness of breath	Associated sensations Preferred body positions Modalities	e.g., tightness of the chest e.g., must sit straight
	Palpitation		
	Joint or muscle pain	Nature of muscle pain during fever Bone pain Back pain	e.g., generalized, or only in limbs
	Headache	Nature and location of the pain Modalities	e.g., like a band around head e.g., better after profuse urination
	Intestinal symptoms, such as diarrhea or vomiting	Description of vomiting Description of diarrhea Other stomach symptoms Modalities	e.g., feels like water stays in the stomach after drinking e.g., pain or burning
	Nasal symptoms	Loss of smell Loss of taste	
	Eye symptoms, such as conjunctivitis		
Unstable gait			
<i>Mental/energy status</i>	Weakness, confusion, fears, restlessness, etc.		e.g., difficulty opening the eyelids
<i>Thirst status</i>	Thirst status in relation to the fever Types of drinks desired Drink-related modalities		e.g., thirstless during fever e.g., hot or cold drinks e.g., desire to take small sips
<i>Weather/temperature</i>	How the weather affects the symptoms		e.g., worse in heat
<i>Time modalities</i>	Worse at particular times of the day		
<i>Perspiration</i>	Nature of perspiration Odor of perspiration Modalities		e.g., oily e.g., offensive e.g., better after sweating
<i>Miscellaneous symptoms</i>	Redness of the face Cold sores Tongue coating		

Apart from these symptoms, demographic characteristics such as age and sex were recorded, as well as (if available) the following COVID specific baseline data: date when symptoms started; results of any PCR and/or antibody tests; CT (computed tomography) status; need for oxygen and/or ICU care, if hospitalized. The full questionnaire is available as **►Supplementary file 1** (available online only). After screening of patients by a study team member, the questionnaire was administered to consenting eligible individuals.

Four members of the Hong Kong team entered the data collected into the Clificol database. Some further information, such as the use of traditional Chinese medicines for the COVID symptoms, was extracted from the patients' notes or, if not recorded, by retrospectively asking the patients.

The data were collected in the online Clificol database, which is a Cloud-based, fully GDPR/HIPAA compliant international clinical case registry (<https://www.clificol.net/>). The analysis team could download the data periodically from the platform as Excel sheets. To ensure data integrity, we checked that all cases had linked consultations and repertory data. Any errors detected were corrected by discussing with the Clificol database team.

We defined a "prescription episode" by grouping together a given consultation where a remedy was prescribed with any subsequent consultations in which the prescribed remedy remained unchanged. Therefore, a prescription episode consists of one or more consultations where the remedy remained unchanged. This enabled us to assess the evolution of the patient until a change in prescription and/or the discharge of the patient. Each case could therefore consist of more than one prescription episode.

The data analysis consisted of various descriptive analyses. In addition to this, we wanted to find out if a distinct symptom picture, or symptom pictures, could be identified in the data. For this purpose, we made use of principal component analysis (PCA) and cluster analysis. PCA is a commonly used statistical method to extract the most meaningful dimensions in a large dataset. For example, in a dataset of symptoms and characteristics of patients including those with a cold and some without that illness, PCA would typically show that having versus not-having a runny nose is a meaningful descriptor whereas the height of the patient would not come out as a particularly useful descriptor of the dataset. Under the assumption that there is a single COVID-19 genus epidemicus, one would expect to see one single cluster of symptoms with patient-dependent variations around one core set of symptoms/rubrics.

To identify the (number of) cluster(s) that best fits the observed data, the *k*-means algorithm was used.<sup>26</sup> The *k*-means algorithm separates the data into a pre-specified number of clusters, in a way that minimizes the distance of the data points (=rubrics) to the center of the clusters. Different numbers of clusters can be hypothesized and scored (using the "silhouette score" approach) to identify how well different numbers of clusters fit the underlying data. The latter helps with the determination of the optimal number of clusters for a given dataset.

The descriptive analyses were conducted in SPSS (version 27). The PCA and *k*-means procedures were performed using Python (version 3.8, [www.python.org](http://www.python.org)) and the scikit machine-learning library (<https://scikit-learn.org/>).

RadarOpus (version 2.2.16) was used for the repertorisations of the symptoms.

## Results

The pandemic was mostly confined to a single wave in early 2020, with only few cases after. The recruitment and selection process of the prescription episodes that were analyzed are given in **►Fig. 1**.

The 359 eligible cases with at least two consultations (766 consultations in total) yielded 363 prescription episodes, of which 357 had homeopathic rubrics available for analysis. Some demographic and clinical characteristics are given in **►Table 2**.

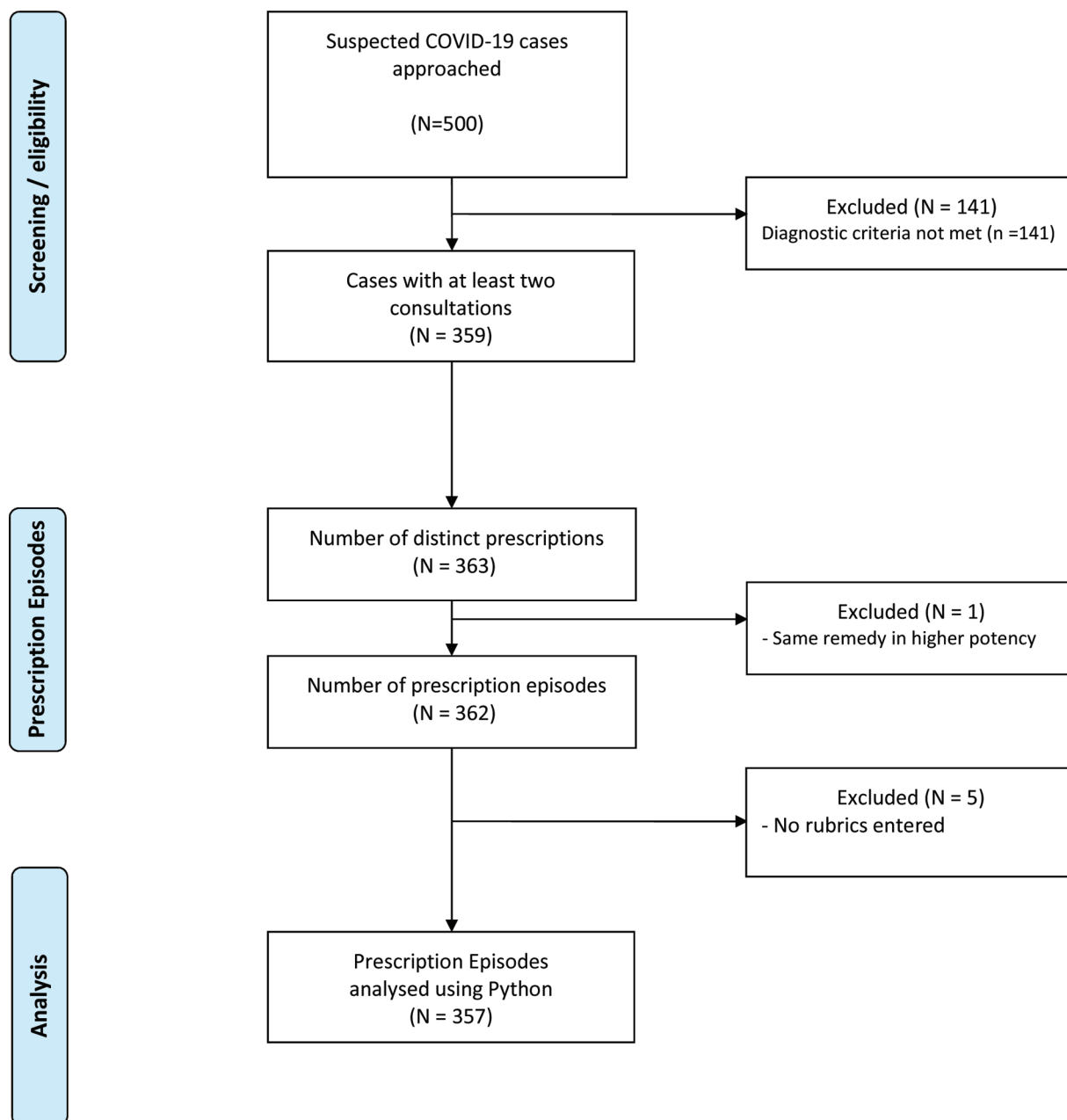
**►Table 2** indicates that there were slightly more female than male patients, and that 75% of the individuals were younger than 45 years of age. It is noteworthy that almost 69% of the patients were seen more than 15 days after the onset of the disease. Almost all of the cases had laboratory confirmed COVID-19, the great majority of which were of mild severity. *Gelsemium sempervirens* was by far the most commonly prescribed remedy, followed by *Bryonia alba*. Both remedies together covered more than 96% of all the prescriptions. *Bryonia alba* was more often prescribed than *Gelsemium sempervirens* in patients that sought homeopathic treatment more than 30 days after the onset of disease (32 vs. 22%). Ninety-five percent of patients had two consultations, and 99% of patients received one prescription only.

### Analysis of Symptoms/Rubrics

In total, 186 distinct rubrics were used. This is more than 150 items in the questionnaire. The reason for this difference is that some rubrics consisting of a combination of individual items in the questionnaire were entered separately. Moreover, there was a possibility to add further symptoms/rubrics in the questionnaire as free text. The majority of these distinct rubrics were used only once or twice. In all the consultations taken together, a total of 8,296 rubric entries were recorded, amounting to a median of 22 rubrics (mean = 23.2) per prescription episode.

In the first phase of the symptom analysis, we conducted a PCA restricted to the rubrics that were reported more than 10 times, which reduced the total number of rubrics included to 60. This PCA yielded two distinct clusters of symptoms (**►Fig. 2**).

Each dot in **►Fig. 2** represents a prescription with its own unique set of associated rubrics. The PCA as a data reduction technique enables us to see that these prescription-associated rubric clusters are not just randomly distributed, but in fact concentrated in at least two clusters. It should be clarified that the PCA separates out clusters 1 and 2 based solely on the symptoms associated with each prescription, so *no* information about the remedy that was prescribed is used.



**Fig. 1** Flowchart of patient recruitment and prescription episode selection.

In the second phase of the rubric analysis, we superimposed the homeopathic remedy prescribed on each of the dots. This revealed that the largest cluster corresponds to cases in which almost exclusively *Gelsemium sempervirens* (blue dots) was prescribed (in 284 of 297 prescription episodes) and a second cluster in which almost exclusively *Bryonia alba* (green dots) was prescribed (in 58 of 61 prescription episodes). Thus **Fig. 2** shows how the PCA algorithm separates the prescriptions according to the set of symptoms reported, irrespective of the remedy prescribed. The prescriptions form two very distinct clusters, and the depth of the shading indicates how many prescriptions there are in a given area. The superimposition in color of the remedies prescribed shows that the PCA prescription clusters corresponded very well to the prescriptions of *Gelse-*

*mium sempervirens* (blue dots) and *Bryonia alba* (green dots). We emphasize again that the PCA-generated clustering of the prescriptions (the dots in the figure) was based purely on the symptoms associated with the prescriptions.

The separation between the two symptom clusters is quite striking if one considers that the dataset consisted of answers given independently by patients to a questionnaire of COVID-related symptoms. We conducted a “Normality” test (Shapiro-Wilk) to assess how compatible these findings were with the hypothesis of a single, Normally distributed cluster of symptoms. This test yielded a  $p$ -value of  $10^{-29}$ , enabling us to very confidently reject the “single cluster” hypothesis. **Fig. 2** also suggests that there may be more than two clusters, possibly related to *Arsenicum album* (red dots) and *Phosphorus* (purple dots) prescriptions. The

**Table 2** Main demographic and clinical characteristics of the patients

Characteristics	Female (n = 199)	Male (n = 160)	Total (N = 359)
Age [years] (mean [min, max])	37.7 [19, 68]	43.1 [22, 66]	40.9 [19, 68]
15–29 (percentage of column total)*	25 (13%)	13 (8%)	38 (11%)
30–44	139 (70%)	89 (56%)	228 (64%)
45–59	32 (16%)	43 (27%)	75 (21%)
60–74	3 (2%)	15 (9%)	18 (5%)
Duration of symptoms; n (%)			
2–6 days	3 (2%)	7 (2%)	10 (3%)
7–9 days	14 (7%)	21 (13%)	35 (10%)
10–14 days	38 (19%)	29 (18%)	67 (19%)
15–30 days	95 (48%)	64 (40%)	159 (44%)
> 30 days	49 (25%)	39 (24%)	88 (25%)
Accuracy COVID-19 diagnosis; n (%)			
Suspected	9 (5%)	3 (2%)	12 (3%)
PCR / Ab / Ag confirmed	190 (96%)	157 (98%)	347 (97%)
Severity of COVID-19; n (%)			
Mild	175 (88%)	143 (89%)	318 (89%)
Moderate	24 (12%)	17 (11%)	41 (11%)
Remedy prescriptions; n (%)			
<i>Arsenicum album</i>	2 (1%)	2 (1%)	4 (1%)
<i>Bryonia alba</i>	41 (20%)	24 (15%)	65 (18%)
<i>Capsicum annuum</i>	1 (1%)	0 (0%)	1 (0%)
<i>Eupatorium perfoliatum</i>	0 (0%)	1 (1%)	1 (0%)
<i>Gelsemium sempervirens</i>	153 (76%)	132 (82%)	285 (79%)
<i>Phosphorus</i>	4 (2%)	3 (2%)	7 (2%)

\*Percentages rounded to the nearest integer.

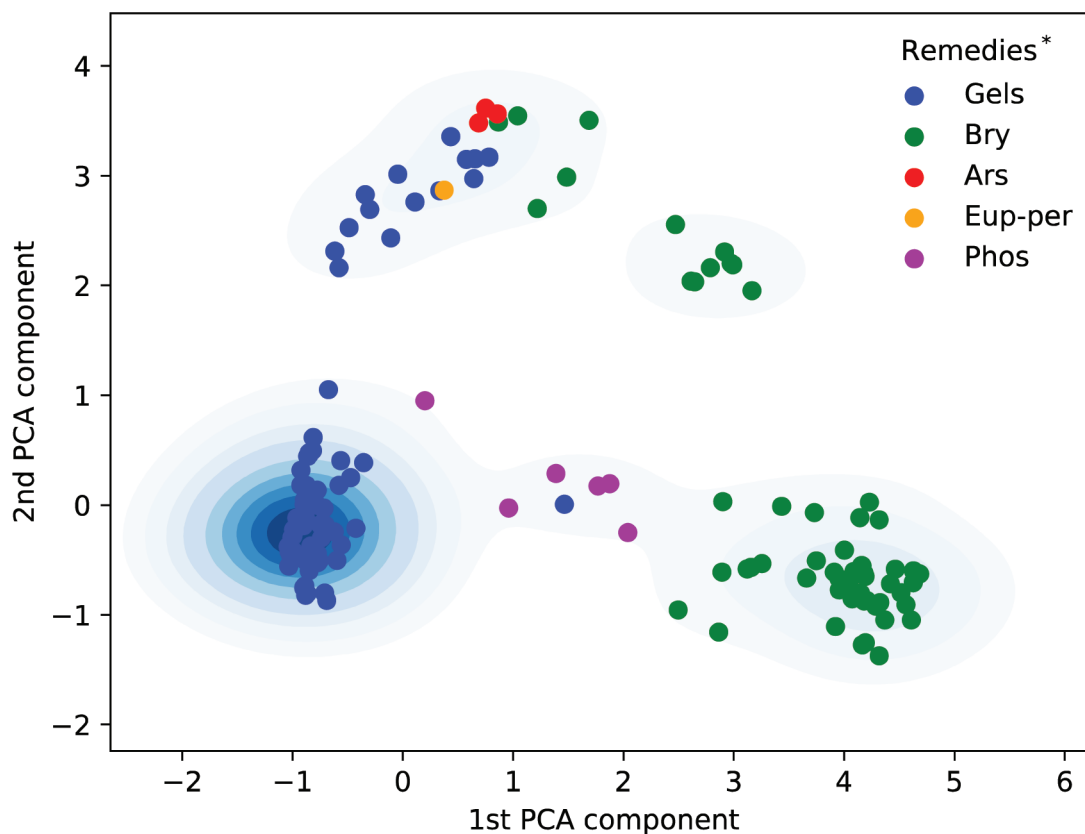
numbers of these cases are, however, too low to state this confidently.

In the third phase of the rubric analysis, we attempted to determine the optimum number of clusters via an automatic clustering of the data using the *k*-means algorithm. This analysis identified the optimal number of clusters to be 2, which is consistent with the *Gelsemium sempervirens* and *Bryonia alba* clusters as identified by the PCA. The hypothetical *Arsenicum album* and *Phosphorus* clusters were too small to be picked up by the automatic clustering algorithm.

In the fourth phase of the rubric analysis, we looked more closely at the specific nature of the rubrics, including in relation to the identified clusters. This analysis identified a set of 13 very common symptoms with a prevalence ranging from 82 to 97% in *both* the identified clusters, and therefore effectively in all the cases. These rubrics are, in descending order of prevalence: GENERALS—SLOW MANIFESTATION; GENERALS—PERSPIRATION—after—amel.; BLADDER—URINATION—involuntary—cough agg., during; FEVER—ALTER-NATING WITH—chills; GENERALS—WEATHER—wet weather—agg.; SLEEP—SLEEPINESS—influenza, during; EX-

PECTORATION—INFREQUENT; COUGH—EXHAUSTING; COUGH—TICKLING—larynx, in; SLEEP—SLEEPINESS—cough —with; HEAD—CONSTRICTION—band or hoop; MIND—DULLNESS—chill, during; MIND—DULLNESS—heat—during.

Since these symptoms were unhelpful from a homeopathic perspective in terms of individualizing the remedy choice, we conducted further analyses to identify the rubrics that best differentiated the two identified symptom clusters. For this, we identified the rubrics that best separated the two clusters on the basis of the PCA, which ascribes a score to each symptom that is positive for symptoms most representative for cluster 1, and negative for symptoms most representative for cluster 2. A score of zero indicates that the symptom is not representative for either cluster. Rubrics with a PCA score that deviated >10% (>0.1 for cluster 1, < -0.1 for cluster 2) from zero were selected. These analyses yielded six differentiating rubrics for cluster 1 and 25 differentiating rubrics for cluster 2. The top bar chart in **Fig. 3** depicts the six rubrics sorted in descending order of prevalence in cluster 1, and the bottom bar chart depicts the 25 rubrics in descending order of prevalence in cluster 2.



**Fig. 2** Distribution of prescriptions along the first two PCA components, with the remedies prescribed superimposed in color. The depth of the shading around the clusters is proportional to the number of prescriptions.

Repertorisations of the symptoms in the *Gelsemium sempervirens* and *Bryonia alba* clusters are given in ►**Supplementary files 2 and 3**, respectively (available online only). These repertorisations further confirm the usefulness of these symptoms in linking the appropriate homeopathic remedy to the recorded symptoms/rubrics.

The repertorisation of the 21 most commonly observed symptoms is given in ►**Supplementary file 4** (available online only). This table indicates that relying simply on the totality of the most common symptoms could have led to the less common cluster of *Bryonia alba* symptoms being missed out. This suggests that modern statistical techniques can play a complementary role in identifying genus epidemicus symptom clusters.

## Discussion

We found that there were at least two symptomatic clusters that should be acknowledged as the COVID-19 genus epidemicus of the Chinese population. Moreover, these symptom clusters were very strongly associated with the *Gelsemium sempervirens* and *Bryonia alba* remedy pictures.

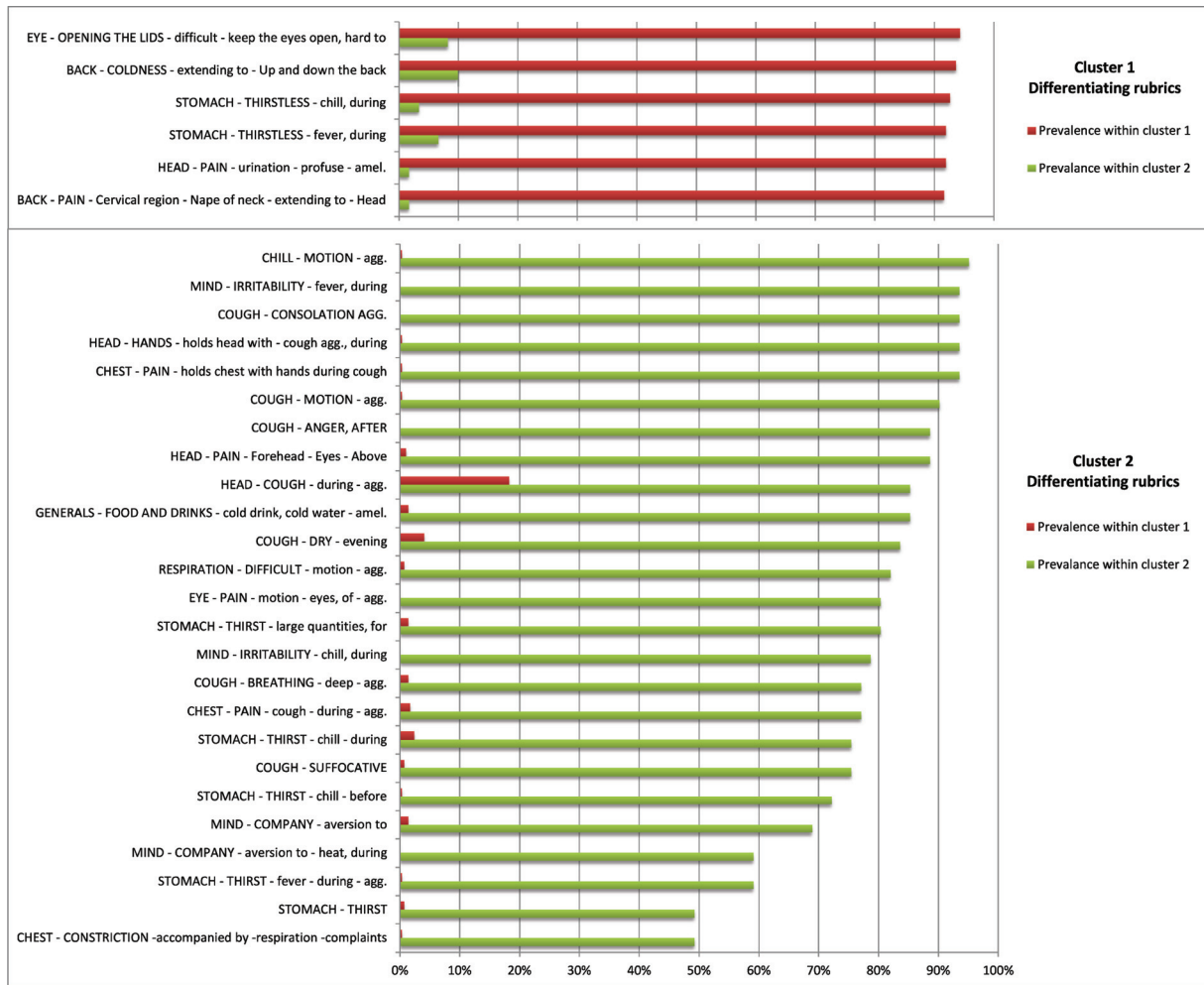
The main strength of this work is that for the first time we made use of modern and powerful statistical analysis techniques to identify symptom clusters observed in a large group of patients during an epidemic. It was also noteworthy that these—purely symptomatically defined—clusters were

highly associated with two homeopathic remedy pictures. In other words, these data confirm the observation by homeopaths that patients present distinct remedy pictures, which can be observed during epidemics.

Further strengths were that the study population was relatively homogeneous, and that the symptoms were obtained and identified in a standardized way with the help of a questionnaire. Both factors contributed to reduced “noise” (variability) in the data, and thereby enhanced the ability to detect the “signal” of the genus epidemicus, via an improved “signal/noise” ratio.

A weakness inherent in any case registry is that these analyses primarily generate, but do *not* confirm, any specific hypotheses. We therefore considered alternative explanations for the observation of at least two distinct genus epidemicus symptom clusters. One consideration is that specific biases could have affected the selection of the symptoms and rubrics. For instance, confirmation bias in the selection of symptoms by homeopaths is well known, and has been referred to as a possibility in the selection of retrospectively collected COVID case materials.<sup>18</sup> However, use was made of a pre-specified patient administered questionnaire to obtain the great majority of the symptomatic data. This likely reduced bias, by making the symptom selection less dependent on the preferences and biases of individual homeopaths.

Whilst the use of a questionnaire is a strength in some ways, it could also be considered as a weakness in other ways.



**Fig. 3** Repertory rubrics of symptoms that best differentiate cluster 1 (top chart) and cluster 2 (bottom chart), in descending order of their prevalence within respectively cluster 1 and 2.

It could be argued that a questionnaire is likely to artificially inflate the number of rubrics via a lowering of the “symptom entry threshold”. The latter appears to be borne out by the data: the median of 22 rubrics used per prescription episode in the Chinese population is much higher than the median of four to five rubrics per prescription episode in the non-Chinese patients. In some ways this difference is unsurprising, because the Chinese rubrics were *extracted* from a 150-item questionnaire, while the rubrics from the rest of the world were *selected* by the homeopath. Nonetheless, many rubrics that were commonly observed in the Chinese population were used far less, or not at all, in the COVID-19 cases from the rest of the world. This included certain keynote symptoms such as “headache ameliorated by profuse urination” that, if observed, would probably have been entered in the repertorisation by homeopaths from the rest of the world. It is therefore difficult to fully explain some of the big differences between the high prevalence of certain rubrics in the Chinese population and the low prevalence of the same rubrics in the rest of the world. It is of course a possibility that different geographical populations present distinct symptom pictures. These questions need to be further explored, and we aim to do so in a subsequent paper

that further analyses the symptom rubrics used in COVID-19 cases from the rest of the world.

Overall, our study suggests that the use of a questionnaire was useful, in particular because it reduced variability in the rubrics entered. We feel that for the purposes of identifying genus epidemicus symptom clusters, the advantages related to the standardized nature of a questionnaire outweigh the disadvantages of potentially missing out on relevant symptoms and/or over-reporting of symptoms.

The PCA and *k*-means cluster analyses are modern techniques to identify patterns in large amounts of data. Differently put, these techniques enable the identification of patterns in data that are characterized by a non-random distribution.

A potential bias affecting the *k*-means algorithm is that the number of clusters needs to be specified prior to the actual analysis. We therefore ran and scored multiple analyses with a pre-specified number of clusters ranging from 2 to 10. This identified two clusters as the best fit (slightly better than three clusters). This corresponds well to what we see in **Fig. 2**. The *k*-means algorithm is most often applied to continuous numerical data, and we applied it to categorical data (presence or absence of symptom). There is no generally agreed way to cluster categorical data, but nonetheless *k*-



means is commonly used for this purpose. Another potential problem is that the *k*-means algorithm finds it difficult to distinguish clusters when the number of dimensions in the data is large and the overlap between members of a cluster is low. Fortunately the use of questionnaires kept the number of dimensions in our dataset low (60 symptoms used more than 10 times) and the overlap high (22 symptoms per prescription on average).

In the Chinese dataset, this analysis identified a *frequently* observed cluster of a relatively *small number of symptoms* that homeopaths would clearly identify as a *Gelsemium sempervirens* symptom picture (→ **Supplementary file 2**, available online only). On the other hand, our analysis also identified a *less frequently* observed cluster consisting of a *larger number of symptoms* identifiable as a *Bryonia alba* symptom picture (→ **Supplementary file 3**, available online only). These findings suggest that only looking at the totality of common symptoms can lead to less frequently observed clusters of symptoms being missed (→ **Supplementary file 4** available online only). Interestingly, even though the distinguishing symptoms of the less frequently observed cluster were far less common, the homeopaths were still able to identify it as a *Bryonia alba* remedy picture in the great majority of cases. Perhaps this was facilitated by the human brain's ability to identify and differentiate patterns in large amounts of data: in this instance, this involved a larger cluster of *Bryonia alba* symptoms. At the same time, it needs to be pointed out that the ability of the clinician's brain to "see" certain patterns in data can be affected by several cognitive biases,<sup>27</sup> including in particular "confirmation bias" (selectively focusing on the data that confirm a pre-conceived notion).<sup>28</sup> Therefore, statistical analyses of data from patient registries such as Clifcol have an important complementary role to play.

The idea of a genus epidemicus remedy in epidemics has been historically focused on finding a single remedy for *prophylactic* use. On the other hand, Hahnemann himself acknowledged that, in the treatment of affected patients, different remedies may be indicated for treating different stages of the epidemic disease.

Apart from the large variety in genus epidemicus remedies proposed for COVID-19, most prospective and retrospective studies investigating the homeopathic treatment of COVID-19 suggest that a multitude of remedies are prescribed.<sup>13,15,19,29</sup> As such, the suggestion that there can be more than one genus epidemicus remedy is not new; James Tyler Kent already suggested that there can be a set of individual remedies related to a particular epidemic.<sup>30</sup> Whilst Hahnemann argued that in healthy individuals a single genus epidemicus remedy was likely to be identifiable, he also suggested in §241 of his *Organon* (6th edition) that pre-existing chronic disease could possibly influence the symptomatic expression of the epidemic disease. The possibility that this could occur in the context of the COVID-19 pandemic has also been raised by others.<sup>31</sup>

We would also like to point out that there is some confusion with the more ancient term "*Genius Epidemicus*" which is defined as "The influence, atmospheric, telluric, or cosmic, or the combination of any two or three, regarded by

the ancients as the cause of epidemic and endemic diseases" (<https://medical-dictionary.thefreedictionary.com/genius-+epidemicus>). Some of the publications are using this outdated term, when they are actually referring to the genus epidemicus remedies.<sup>32,33</sup>

Our analyses found that there were distinct clusters of *Gelsemium sempervirens* and *Bryonia alba* patients. The separation of symptoms corresponding to the remedy pictures was striking. Since these symptom clusters were generated on a purely "data led" basis by a machine learning (artificial intelligence) algorithm, our study provides some further evidence for the homeopathic concept of individualization based on the "remedy picture" as observed in individual patients.

The use of clinical case registries in homeopathy has a long history,<sup>34</sup> and this includes attempts to systematically link outcome to the homeopathic medicines prescribed.<sup>35</sup> The use of a clinical case registry such as Clifcol in conjunction with modern Cloud-based technology, as well as powerful analytical tools, offers unprecedented possibilities to enhance learning from the experiences of homeopaths worldwide. We have illustrated that it is possible to shed new light on an "old" topic such as the genus epidemicus. For the further development of homeopathy, it is necessary to place more emphasis on a "data-based" approach. Apart from answering some questions, this will at the same time challenge existing opinions and also raise new questions. In our opinion, this heralds a new phase in the further development of evidence-based homeopathy.

#### Highlights

- The Clifcol COVID-19 Clinical Case Registry is a database that can be used to address important questions in homeopathy based on "real world" data.
- The genus epidemicus concept was investigated on the basis of symptoms reported by 359 Chinese patients during the first wave of the pandemic.
- Modern techniques, such as principal component analysis and *k*-means clustering, were applied for the first time to reported symptoms and repertory rubrics.
- Patients presented two distinct clusters of symptoms, which had a high overlap with the remedy pictures of the most commonly prescribed remedies: *Gelsemium sempervirens* and *Bryonia alba*.
- Our findings are *not* compatible with the hypothesis of a single COVID-19 genus epidemicus remedy.

#### Supplementary material

**Supplementary File 1.** Full COVID-19 symptom questionnaire.

**Supplementary file 2.** Repertorisation differentiating symptoms cluster 1.

**Supplementary file 3.** Repertorisation differentiating symptoms cluster 2.

**Supplementary file 4.** Repertorisation of the most common symptoms.

### Authors' Contributions

Alexander T. and Yvonne F. were involved in the planning, conduct, analysis, and writing of the article. Robbert van H. and Aaron T. were involved in the planning and interpretation of the analyses and writing of the article. All authors agree to the contents of the paper.

### Conflict of Interest

None declared.

### Acknowledgements

The authors would like to acknowledge Lex Rutten for his critical, constructive and practical feedback, which helped us to improve this article.

### References

- Leung K, Wu JT, Liu D, Leung GM. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *Lancet* 2020;395:1382–1393
- To KLA, Fok YYY. Homeopathic clinical features of 18 patients in COVID-19 outbreaks in Hong Kong. *Homeopathy* 2020;109:146–162
- Malan H. Belladonna versus Scarlatina. *Lancet* 1848;52:330
- Leary B. Cholera and homœopathy in the nineteenth century. *Br Homeopath J* 1987;76:190–194
- Jahn S. Homeopathy in the 1918 influenza pandemic. *Am J Hom Med* 2015;108:160–170
- Bracho G, Varela E, Fernández R, et al. Large-scale application of highly-diluted bacteria for leptospirosis epidemic control. *Homeopathy* 2010;99:156–166
- Nayak D, Chadha V, Jain S, et al. Effect of adjuvant homeopathy with usual care in management of thrombocytopenia due to dengue: a comparative cohort study. *Homeopathy* 2019;108:150–157
- Devos P, Serneels P, Serneels P, Goyens M, Wassenhoven VanM. Survey COVID-19 first peak 2020—viewpoint of Belgian homeopathic medical doctors. *Int J High Dilution Res* 2021;20:51–74
- Devos P, Van Wassenhoven M, Serneels P, Serneels P, Goyens M. Survey COVID-19 first peak 2020. *Int J High Dilution Res* 2021;20:5
- Rao KD, Kaur J, Peters MA, Kumar N, Nanda P. Pandemic response in pluralistic health systems: a cross-sectional study of COVID-19 knowledge and practices among informal and formal primary care providers in Bihar, India. *BMJ Open* 2021;11:e047334
- Waisse S. Severe acute thromboinflammation: case report of individualized homeopathic treatment. *Homeopathy* 2021;110:132–136
- Takacs M, Frass M, Walder G, et al. Adjunctive homeopathic treatment of hospitalized COVID-19 patients (COVIHOM): a retrospective case series. *Complement Ther Clin Pract* 2021;44:101415
- Clapers ML, Velat M, Clapers JM, Vidal-Jové J, Mora J. Analysis of homeopathic genus epidemicus for COVID-19 patients in Spain. *Homeopathy* 2021;110:236–243
- Laville P. Étude descriptive rétrospective d'une cohorte de 16 patients suspects de COVID-19. Short survey. *La Revue d'Homéopathie* 2021;12:22–30
- Jethani B, Gupta M, Wadhvani P, et al. Clinical characteristics and remedy profiles of patients with COVID-19: a retrospective cohort study. *Homeopathy* 2021;110:86–93
- Holandino C, Oliveira AP, Vilhena EC, et al. Homeopathy and the COVID-19 pandemic: quasi-experimental observational study. *Int J High Dilution Res* 2021;20:9–10
- Varanasi R, Nayak D, Khurana A. Clinical repurposing of medicines is intrinsic to homeopathy: research initiatives on COVID-19 in India. *Homeopathy* 2021;110:198–205
- Rutten L, Smedley T, Ives G, et al. Data collection during the COVID-19 pandemic: learning from experience, resulting in a Bayesian repertory. *Homeopathy* 2021;110:94–101
- Manchanda RK, Miglani A, Gupta M, et al. Homeopathic remedies in COVID-19: prognostic factor research. *Homeopathy* 2021;110:160–167
- Adler UC, Adler MS, Hotta LM, et al. Homeopathy for COVID-19 in primary care: a structured summary of a study protocol for a randomized controlled trial. *Trials* 2021;22:109
- To KLA, Fok YYY. Updates from China: clinical and research experience of homeopathy in the pandemic. *Int J High Dilution Res* 2021;20:8
- Tournier AL, Fok YYY. Preliminary analysis of the Clifical COVID-19 support project. *Int J High Dilution Res* 2021;20:6–7
- Jacobs J. Homeopathic prevention and management of epidemic diseases. *Homeopathy* 2018;107:157–160
- Seiler H. The Development of Samuel Hahnemann's Practice as a Physician [Die Entwicklung von Samuel Hahnemanns ärztlicher Praxis]. Haug Verlag; 1988
- Diagnosis and Treatment Protocol for COVID-19. (Trial Version 7), National Health Commission of the People's Republic of China. Accessed April 8, 2020 at: <http://www.nhc.gov.cn/xcs/zhengcwj/202003/46c9294a7dfe4cef80dc7f5912eb1989.shtml>
- Arthur D, Vassilivskii S. K-means ++: the advantages of careful seeding. *Proceedings of the Annual ACM-SIAM Symposium on Discrete Algorithms* 2007:1027–1035
- Norman GR, Monteiro SD, Sherbino J, Ilgen JS, Schmidt HG, Mamede S. The causes of errors in clinical reasoning: cognitive biases, knowledge deficits, and dual process thinking. *Acad Med* 2017;92:23–30
- Croskerry P, Norman G. Overconfidence in clinical decision making. *Am J Med* 2008;121:S24–S29
- Manchanda RK, Miglani A, Chakraborty M, et al. Impact of bias in data collection of COVID-19 cases. *Homeopathy* 2022;111:57–65
- Kent JT. *Lectures on Homeopathic Philosophy*, Lecture 3. B. Jain Publishers; 2003
- da Costa Fujino FMS, Olandim AACC, Barnabé VD, Coggan JA, Benites NR. A homeopathic view of the influence of chronic diseases in the manifestation of COVID-19. *Homeopathy* 2021;110:67–69
- Bala R, Srivastava A. Historical journey of homoeopathy during epidemic diseases in the light of 2019 novel coronavirus pandemic. *Int J Sci Healthc Res* 2020;5:215–233
- Gosik MS, Mendes MFX, Werneck Dos Santos LMA, et al. Medicines for the new coronavirus in the view of Classical Systemic Homeopathy. *Complement Ther Clin Pract* 2021;45:101482
- Van Haselen RA, Fisher P. Analysing homœopathic prescribing using the READ Classification and information technology. *Br Homeopath J* 1990;79:74–81
- Van Haselen RA, Fisher P. Towards a new method for improving clinical homœopathy. Correlating homœopathic prescription with outcome in rheumatoid arthritis using standard clinical assessments and information technology. *Br Homeopath J* 1992;81:120–126