

Multivariate analysis of homeopathic prescriptions

Newsletter Likelihood Ratio project October 2007
Lex Rutten, Erik Stolper, Roland Lugten, Rob Barthels

The Likelihood Ratio (LR) project assesses the clinical value of six homeopathic symptoms. The prospective assessment the symptoms . 'Diarrhoea from anticipation', 'Fear of death', 'Grinding teeth during sleep', 'Recurrent herpes lips', 'Sensitive to injustice' and 'Loquacity' started June 2004.

February 2007, 3367 patients entered the study and 3246 prescriptions were evaluated. Results were shown in our newsletter of May 2007. In this news letter we mentioned the possibility of correlation of some symptoms for some medicines. After May we explored this subject further. Does correlation influence the LR value of the combination of symptoms? Correlation matrices can detect possible correlations between two symptoms. We tested 300 possible correlations for the 20 most prescribed medicines. Correlations between more than two symptoms can be tested by Principal Components Analysis (PCA).

Another possibility of multivariate analysis is Canonical Discriminant Analysis (CDA). CDA is a statistical analysis, rendering 'types' of patients. The maximum number of types is restricted by the number of assessed symptoms (maximum 6 types) and the number of medicines (number of medicines minus one). These analyses were performed using the statistical package SPSS.

Computers enable us to analyse huge amounts of data. One of the possibilities is pattern recognition.

	Medicine	frequency	percent
1	argentum nitricum	24	2,7
2	arsenicum album	25	2,8
3	calcareo carbonica	64	7,2
4	calcareo phosphorica	23	2,6
5	carcinosinum	35	3,9
6	causticum	38	4,3
7	graphites	25	2,8
8	ignatia	23	2,6
9	lachesis	30	3,4
10	lycopodium	72	8,1
11	mercurius	44	4,9
12	natrium muriaticum	129	14,5
13	nux vomica	30	3,4
14	phosphoricum acidum	19	2,1
15	phosphorus	55	6,2
16	pulsatilla	55	6,2
17	sepia	76	8,5
18	silicea	27	3,0
19	staphisagria	25	2,8
20	sulphur	72	8,1
	Total	891	100,0

Table 1: The 20 most prescribed medicines in the LR project

An example: Satellite pictures consist of pixels in various densities in various parts of the colour-spectrum. These densities vary according to the kinds of soil or vegetation. The pixel-density of each vegetation has a certain mean and a certain intra-group variance specific for that vegetation. Another vegetation has another mean pixel-density and the difference between the pixel-densities of two vegetations is called the inter-group variance. By calculating the difference between intra-group variance and inter-group variance of parts of the picture the computer can tell which kind of vegetation is in a certain part of the satellite image. This is done by CDA. In homeopathy, a specific symptom has different 'densities' in various populations cured by different medicines. We can use CDA to show which symptoms correlate best with which medicines.

The LR project was not designed to perform CDA, but we can get an impression of the possibilities of CDA analysing the prescriptions of 10 medicines that were frequently prescribed and seem to have a relation with the six investigated symptoms.

Principal Component Analysis (PCA) is used to discover inter-relatedness of symptoms. An example: If we compare abilities of schoolchildren, like geometry, arithmetic and algebra, we will see much

correlation because these are all mathematical abilities. Three abilities can be summarised in one denominator. With PCA we tested possible relations between the six symptoms in this assessment.

Results of multivariate analysis

There were about 1700 prescriptions with the desired result (GHHOS 2-4), 20 medicines were responsible for 891 of these desired results, see table 1. For these 20 medicines we checked correlation between two symptoms using correlation matrices, and correlation between all symptoms using Principal Component Analysis (PCA). The discrimination between medicines can be assessed by Canonical Discriminant Analysis (CDA). The number of possible discriminant functions is limited by the number of symptoms. We performed CDA for 10 medicines that seem related to the investigated symptoms.

Correlation matrices

We investigated 300 possible correlations and found hardly any correlations between two symptoms. Those correlations are in Table 2. For these medicines we calculated the difference between the combined symptoms and the single symptoms.

medicine	symptoms	LR symptom	LR1*LR2	LR(1+2)	95% CI LR(1+2)
Ignatia	1. fear death	1. 1.73	1.76	5.39	1.40 to 20.78
	2. herpes	2. 1.02			
Nux-v	1. diarrhoea	1. 1.50	1.17	3.55	1.19 to 10.58
	2. herpes	2. 0.78			
ph-ac	1. herpes	1. 1.23	1.73	2.19	0.59 to 8.19
	2. injustice	2. 1.41			
phosphor	1. injustice	1. 0.99	0.92	2.25	1.17 to 4.33
	2. loquacity	2. 0.93			

Table 2: The difference between combined LR for two symptoms after measurement of separate symptoms (LR1*LR2) and after measurement of the combined symptom (LR(1+2)).

The LR of the combined symptoms is calculated by dividing the prevalence of both symptoms together in the medicine population by the prevalence of both symptoms together in the rest-population. We see that the combination of 'Fear of death' and 'Herpes lips' is a stronger indication (LR=5.39) for *Ignatia* than expected from the LRs calculated for the separate symptoms. If both symptoms were independent their combined LR would be $1.73*1.02=1.76$.

Likewise the combination of 'Diarrhoea from anticipation' and 'Herpes lips' has a higher LR for *Nux vomica*. The combination of 'Sensitivity to injustice' and 'Loquacity' is an unexpected indication for *Phosphor*, as the separate symptoms both do not indicate this medicine. The symptoms 'Herpes lips' and 'Sensitivity to injustice' do not significantly increase the indication for *Phosphoricum acidum*, despite their moderate correlation (correlation = 0.431).

Results of Principal Components Analysis

PCA was used to check if one or more of the symptoms were redundant, because of inter-relatedness with the other symptoms. This is done by making a 'scree-plot', which indicates the amount of information given by each symptom (see Figure 1). If the 891 patients in our population responding to 20 medicines could have been described by less symptoms we would see an 'elbow' in this plot. The elbow indicates that more components give no extra information. This is not the case, We need six components to retrieve all necessary information. This indicates that subgroups of patients cannot be differentiated by less symptoms.

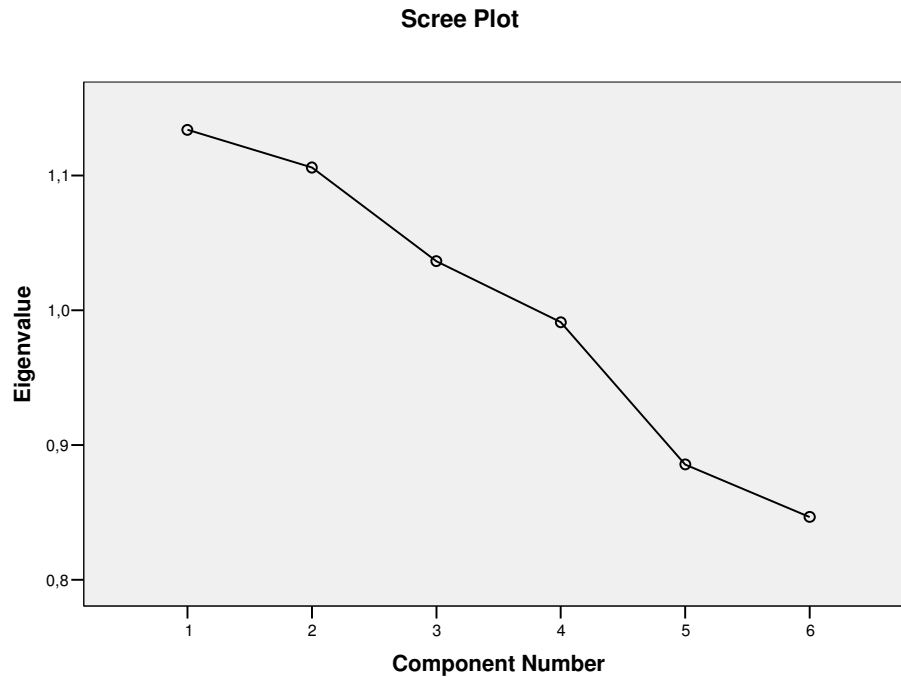


Figure 1: Scree plot indicating the amount of information supplied by each of six symptoms. There is a linear development of information indicating an even spread of information given by each symptom

Results of Canonical Discriminant Analysis

We performed CDA on the population responding to 10 medicines that seem related to the six assessed symptoms. The results of CDA look much like the results of logistic regression. The first discriminant function of this analysis is:

$$\text{Discriminant1} = 0.867 \cdot \text{diarrhoea} + 0.412 \cdot \text{fear} - 0.069 \cdot \text{grinding} - 0.053 \cdot \text{herpes} - 0.353 \cdot \text{injustice} + 0.227 \cdot \text{loquax}$$

Our population consists of ten subgroups, each responding to a different medicine. With six symptoms we can only discern six subgroups. If we divide our population in six groups, there is a group of patients that have similarities in that respect that they have diarrhoea form anticipation and fear of death; they also tend to loquacity. They do not tend to sensitivity to injustice, herpes lips and grinding teeth at night.

The other discriminant functions can be interpreted in a similar fashion, see the following table:

Standardized Canonical Discriminant Function Coefficients						
	Function					
	1	2	3	4	5	6
diarr	,867	-,421	,291	-,106	-,070	,025
fear	,412	,476	-,319	,657	-,060	,282
grind	-,069	-,071	-,037	,087	,867	,499
herpes	-,053	,073	-,181	-,402	-,403	,807
injust	-,353	-,193	,663	,583	-,169	,218
loquax	,227	,676	,614	-,331	,137	,011

We can also analyse which medicines were effective by each profile and which not.

Discriminant function1	
medicine	score
argentum nitricum	1,743
phosphoricum acidum	0,718
lachesis	0,274
arsenicum album	0,27
staphisagria	0,059
sepia	-0,023
carcinosinum	-0,074
natrium muriaticum	-0,303
mercurius	-0,321
causticum	-0,377

The most effective medicine has the highest score. Homeopathic physicians will not be surprised that the medicine *Argentum nitricum* is the most effective medicine for patients out of group 1 because diarrhoea from anticipation and fear of death are the main symptoms. *Causticum* is the least effective of the 10 assessed medicines for this group.

The ordering of medicines that fit the second profile is given in the following table:

Discriminant function 2	
Medicine	score
lachesis	1,022
arsenicum album	0,506
sepia	0,196
carcinosinum	-0,042
staphisagria	-0,086
natrium muriaticum	-0,134
mercurius	-0,155
causticum	-0,19
argentum nitricum	-0,201
phosphoricum acidum	-0,968

In the group belonging to function 2 loquacity is the most common trait, so the effectiveness of *Lachesis* is expected. *Arsenicum album* is also effective in this group because of the fear of death. The limitation of CDA in this case is that we can discriminate only 6 groups because we have only 6 symptoms.

If we do the same for the other discriminant functions we see that *Causticum* works best in function 3 because of the sensitivity to injustice that predominates in this group. *Lachesis* has good results in this group, probably because of the loquacity. We cannot tell if diarrhoea from anticipation is an indication for *Causticum* or *Lachesis*; this group is a combination of patients responding to three medicines (also phosphoricum acidum). See following table.

Discriminant function 3	
Medicine	score
causticum	0,716
lachesis	0,355
phosphoricum acidum	0,188
Mercurius	0,083
Sepia	0,055
Staphisagria	0,043
Carcinosinum	0,01
argentum nitricum	-0,023
natrium muriaticum	-0,277
arsenicum album	-0,577

If we were to choose between these 10 medicines combining these three discriminant functions we would prefer *Argentum nitricum* if the patient had diarrhoea from anticipation and was not sensitive to injustice. If we translate these data into materia medica we could represent this medicine like:

Materia Medica Argentum nitricum

Positive symptoms: **Diarrhoea from anticipation**, ..., ..., etceteras

Negative symptoms: Sensitive to injustice, ..., ..., etceteras

Using discriminant analysis we can be more confident about symptoms that indicate the medicine, but we can also use symptoms that contraindicate it, which is an addition to the existing materia medica.

Discussion

The multivariate analysis of our database had two goals:

1. Detect interrelations between symptoms.
2. Detect patterns of symptoms in our population.

For the first goal we made correlation matrices and performed PCA. We found hardly any correlation between symptoms for the 20 most frequently used medicines. PCA gave no indication that the number of symptoms could be reduced for these 20 medicines.

The LR project was designed to compare symptoms, not medicines. The LR results offer new possibilities. Homeopathic physicians use present as well as absent symptoms, but their instruments (repertory and materia medica) refer only to symptoms that can be expected and only if the symptom is present. LR gives indications about symptoms not to expect when a certain medicine is considered e.g. herpes lips considering *Graphites* ($p=0.154$). Negative LR also gives us the possibility to estimate the meaning of the absence of an expected symptom.

LR assessment is a long-term project. We need to assess more than 500 symptoms to get a rather complete evaluation of the most frequently used homeopathic symptoms. This will improve our method step-by step but takes a considerable amount of time. One of the advantages of LR assessment is that any medicine can turn up to be useful if a particular symptom is present.

Discriminant analysis is suited to maximise differences between different groups (medicine populations) by a combination of present and absent attributes. Discriminant analysis allows us to evaluate a number of medicines at the same time. If we evaluate the 20 most prescribed medicines we will improve half of our cases. CDA will give us information about which symptoms to expect and which symptoms not to expect when we consider a certain medicine. It gives, however, no information about other medicines and we cannot differentiate these 20 medicines from other medicines. CDA does not give the information of negative LRs; the meaning of the absence of an expected symptom is not calculated.

We performed CDA on the populations responding to 10 medicines and evaluated six symptoms, each symptom is considered to be a keynote for different medicines. This is not ideal for this purpose, we should have had more symptoms than medicines. But still we see that the *Argentum nitricum* patient is unlikely to be sensitive to injustice and that the *Causticum* patient is unlikely to have fear of death.

Future possibilities for discriminant analysis

Discriminant analysis is an interesting option to improve homeopathic prescribing. We should consider CDA for evaluation of cured cases considering the twenty most prescribed medicines. Each medicine has a number of 'keynotes', symptoms that occur rather frequently in the population that responds well to that medicine and are the most indicative for that medicine. CDA would be effective if we assess a number of symptoms for each medicine. If we choose, say, five keynote-symptoms for each of 20 medicines this would constitute a questionnaire of 100 symptoms (maximum). If that questionnaire is given to a sufficient number of patients that responded well to the investigated medicines we could get a reliable differential diagnosis of our most successful medicines, including positive and negative indications for each medicine.