Homeopathy (2006) 95, 88–93

doi:10.1016/j.homp.2006.01.001, available online at http://www.sciencedirect.com

ORIGINAL PAPER

A Bayesian perspective on the reliability of homeopathic repertories

ALB Rutten*, CF Stolper, RFG Lugten and RWJM Barthels

Commissie Methode en Validering VHAN (Dutch Association of Homeopathic Physicians), The Netherlands

An interim assessment of likelihood ratio of homeopathic symptoms shows that there are serious flaws in Kent's repertory. The system of expressing relationship between symptoms and expected results from medicines by typeface is unclear and unreliable. Bayesian methods can improve this and expand the possibilities of the repertory including the possibility of interpreting the absence of expected symptoms. *Homeopathy* (2006) **95**, 88–93.

Keywords: likelihood ratio; Bayesian analysis; homeopathy; repertory; clinical symptoms; inclusion; exclusion

Introduction

We recognise, for instance, the patient who will likely benefit from the homeopathic medicine *Rhus toxicodendron* by the fact that his complaints are ameliorated by motion. Experience tells us that this trait is often seen in patients who respond well to this medicine, more often than in patients responding to other medicines. On the other hand, patients who respond well to *Arsenicum album* are infrequently warm-blooded; we will hesitate to prescribe *Arsenicum album* if the patient is not cold-blooded.

Thomas Bayes described the idea of such experience resulting in practical knowledge in 1763. He even produced a mathematical formula to calculate the possibility that a medicine will work based on such experience. Suppose that 'amelioration from motion' is seen 10 times as frequently regarding patients that respond well to *Rhus toxicodendron*. Then the odds (odds = chance/(1-chance)) that *Rhus toxicodendron* will work improve 10 fold if the patient experiences 'amelioration from motion' (posterior odds = LR * prior odds). Bayes' formula is widely accepted and successfully used in all kinds of computer software.

In homeopathy we use repertories showing which medicines are indicated when certain symptoms are present. The most-used repertory is Kent's, which uses different typeface to grade the importance of medicines for each symptom.[1] The meaning of the typefaces in the repertory is not very clear. Kent gives no explanation in the preface to his repertory. We find some explanation in one of the sources of the repertory, Hering's 'Guiding Symptoms'[2]. Hering describes: 'Symptoms occasionally confirmed', 'Symptoms more frequently confirmed', 'Symptoms verified by cures' and 'Symptoms repeatedly verified'. These are absolute, not relative statements, so the more a medicine is prescribed, the more frequently corresponding symptoms will be confirmed and verified.

This system of absolute grading of medicines in rubrics is fundamentally wrong and should be replaced by relative grading: a symptom is an indication for a medicine when it occurs more frequently in the population cured by the medicine than in the rest of the population.[3] The grading should depend on the difference between medicine-population and rest of the population. If, say, the symptom 'fear of death' occurs in 28% of the 'Arsenicum-population' and in 4% of the rest of the population, it occurs seven times as frequently in the Arsenicum-population. In other words: the likelihood of 'fear of death' is 28% in the Arsenicum-population, 4% in the rest of the population: the likelihood ratio (LR) is 7.

Once we accept LR for grading of medicines in symptom rubrics it is possible to calculate probabilities of cure based on available symptoms. The strength of Bayes' method is that we add up several factors

*Correspondence: Lex Rutten, Aard 10, 4813 NN Breda, Netherlands.
E-mail: lexrtn@concepts.nl
Received 3 November 2005; accepted 19 January 2006

Table 1 Increasing chances of cure by adding symptoms and using Bayes' formula

Symptom no.	*Symptom*	*LR+*	*Chances of success (%)*
0			1
1	Fear of death	7	7
2	Symptom 2	7	34
3	Symptom 3	7	78

(symptoms) that are related to our diagnostic process to find the most suitable medicine. A disadvantage of Bayes' formula is that it works easily with odds but is a bit awkward with chances (chance = odds/(odds + 1)). The use of a computer overcomes that problem. Table 1 shows how chances of cure by *Arsenicum* increase when 'fear of death' is present and when subsequently two other symptoms indicating *Arsenicum* with the same strength are added. We assume that without any indicating symptoms chances that *Arsenicum* will work are 1%.

With three good symptoms we have a reasonable certainty of 78% that the medicine will work (Table 1). That three good symptoms suffice to prescribe is well known in homeopathic practice and is sometimes called a 'tripod'. It is not yet certain that a good homeopathic symptom has LR of 6 or more, but there are some indications coming from the evaluation of our LR project after 15 months.

Type and LR

A correct repertory would indicate the difference between medicine-population and rest of the population, expressed as LR. Above we showed that LR = 6 or higher might correspond with bold type entries in the repertory. A homeopathic medicine should be present in the rubric when the population responding to that medicine shows the rubric-symptom clearly more often than the rest of the population, say more than one and a half times more often. Table 2 shows a possible schema for translating LR into type face. Such a schema should be evaluated in due course for correspondence with actual practice.

Methods

Our assessment of the LR of homeopathic symptoms started in June 2004. Eleven doctors record the presence of six symptoms in every new patient. The symptoms are: 'Diarrhoea from anticipation', 'Fear of death', 'Grinding teeth during sleep', 'Herpes lips' and 'Loquacity'. Prescribed medicines are also recorded and each doctor kept track of the results using a specified Glasgow Homeopathic Hospital Outcome Scale (GHHOS).

After 15 months, in September 2005, an intermediate assessment was performed. At that time 1634 patients had been included and 1049 prescriptions evaluated. For calculating LR values we used all outcomes of 2–4 on the GHHOS scale.

Table 2 The relation between typeface and likelihood ratio

Type	*LR*
Plain	1.5–3.0
Italics	3.0–6.0
Bold	More than 6.0

The intermediate values represented by italics are just a rough estimation; further research must indicate optimal values.

Table 3 The symptom 'fear of death' in relation to three homeopathic medicines

Medicine	*Fear of death, n =*	Total population of medicine, *n =*	*LR+*
Anacardium	2	4	13.36
Arsenicum album	2	8	6.66
Calcarea carbonica	2	27	1.95

We use a three-point scale to record symptoms. Many symptoms are not dichotomous (yes or no statements). If such a symptom is strongly present we will suspect corresponding medicines, eg a strong 'fear of the dark' will indicate the possibility of the medicine *Stramonium*. On the other hand, if we suspect *Stramonium* for other reasons, we will accept a moderate kind of 'fear of the dark' as a confirmation for this medicine. Therefore, we distinguish three degrees of presence of a symptom: 0 = absent; 1 = moderate and 2 = strong. Consensus is necessary to specify this for each symptom. As a general principle for this threshold values we take Grade 2 (strong) as an intensity of the symptom that cannot be controlled by will and is spontaneously mentioned. Grade 1 (mild) is the intensity that is less, but confirmed if a medicine indicated by that symptom is considered. Grade 0 (absent if the symptom occurs seldom or never or is not stronger than in the average population (eg desire chocolate). See also 'Excluding medicines', below.

Results

An example of the difference between absolute and relative occurrence of symptoms is our assessment of the symptom 'Fear of death'. Here we have three medicine-populations each of two persons having strong fear of death, but the total populations from which they are drawn are not the same size. See Table 3.

Calculated on the basis of absolute occurrence *Anacardium*, *Arsenicum album* and *Calcarea carbonica* get the same emphasis. In Kent's repertory *Anacardium*

is in plain type and *Arsenicum* and *Calcarea* in bold type. But *Arsenicum* and *Calcarea* are much more frequently prescribed than *Anacardium*. In absolute terms the symptom is equally common for all three medicines but relatively it is much commoner with *Anacardium*. Here we see that LR expressing the relative occurrence is the only right way. *Anacardium* is much more likely to be indicated if the symptom 'fear of death' is present than *Calcarea carbonica*, despite the rubric in the present repertory.

After translating LR into type we can estimate how many patients who respond well to a medicine should have the symptom under investigation. Comparing this with the data from our prospective assessment gives an indication of the correctness of the rubric. We expected that entries of the most common medicines would often be wrong. For six medicines (*Calcarea carbonica*, *Lycopodium*, *Natrum muriaticum*, *Pulsatilla*, *Sepia* and *Sulphur*) we had more than 20 persons with outcome 2–4. We did not include the symptoms 'sensitive to injustice' and 'diarrhoea from anticipation' here because they are not present or incomplete in Kent's repertory (Table 4).

Even with these small numbers we see repertory entries that are not compatible with the assessed facts.

Table 4 Predicted and actual occurrence of four symptoms in the populations responding to the six most prescribed homeopathic medicines

Medicine	*n*	*Type*	*Expected*	*Present*
Fear of death (strong)				
Calc.	27	**Bold**	6	2
Lyc.	22	*Italics*	3	2
Nat-m.	54	Plain	3	0
Puls.	25	*Italics*	3	1
Sep.	33	Plain	2	1
Sulph.	29	Plain	2	0
Grinding teeth during sleep (strong)				
Calc.	27	Plain	2	1
Lyc.	22	—	≤1	1
Nat-m.	54	—	≤3	2
Puls.	25	—	≤1	1
Sep.	33	Plain	3	2
Sulph.	29	—	≤2	1
Herpes of lips (strong)				
Calc.	27	—	≤1	0
Lyc.	22	—	≤1	0
Nat-m.	54	**Bold**	16	7
Puls.	25	—	≤1	1
Sep.	33	**Bold**	11	3
Sulph.	29	Plain	2	2
Loquacity (strong)				
Calc.	27	Plain	3	1
Lyc.	22	—	≤2	2
Nat-m.	54	Plain	5	0
Puls.	25	—	≤2	1
Sep.	33	—	≤3	2
Sulph.	29	Plain	3	1

Chances that the entry of *Calcarea carbonica* in 'Fear of death' is rightfully bold are 0.03 (one tailed value). The right type is probably plain. For *Natrum muriaticum* (plain) in the same rubric $P = 0.035$. *Natrum muriaticum* should not be entered bold in 'Herpes lips' ($P = 0.003$); italics seems more appropriate. The same goes for *Sepia* in the same rubric ($P = 0.004$). Loquacity is not an indication for *Natrum muriaticum* ($P = 0.004$), the presence of this symptom even pleads against this medicine.

Excluding medicines

We demonstrated that probably many entries, especially for frequently used medicines, in the repertory are incorrect. This does not mean, however, that those symptoms are not possible for people responding to those medicines. We will still encounter loquacious *Sulphur*-patients; the symptom loquacity does not indicate *Sulphur*, nor exclude it. Using Bayesian theory we can now be more precise about this matter.

Bayesian theory is about chances, not certainties. We can seldom draw definitive conclusions from one symptom, positive nor negative. So how do we use LR for inclusion or exclusion? Generally we use inclusion and LR+ at the beginning of a consultation when many medicines are possible. In other words: when prior chance of a medicine is low. Exclusion and LR– is better used at a later stage, when we have a limited number of medicines to choose between (Table 5).

Threshold values for symptoms are handled intuitively in current homeopathic practice and should be made more explicit. It is not easy to determine threshold values for vague symptoms. We should try to use the thresholds that we use in practice, but if we must check the symptom consciously during the first consultation the threshold to consider a symptom present tends to be lower. It is therefore important to indicate the prevalence of the symptom in the total population to enable users to estimate validity of these figures in their situation.

An example: sensitivity to injustice

We have attempted to reach consensus on threshold values in our research group. One example of a vague symptom and its prevalence in our research-population is sensitivity to injustice (Table 6).

Grade 1: Sensitive to injustice without apparent influence on behaviour.

Grade 2: Behaviour influenced, eg protest marches; turns off television if images of injustice shown. Prevalence grade 2 in our research-population = 10%. Prevalence of grade 1+2 in our research population = 27%.

Table 5 Positive and negative likelihood ratios

Positive LR (LR+) is used to calculate changes in odds (or chances) if the symptom is present. Odds become greater when LR+ > 1, higher is better to include.
Negative LR (LR–) is used to calculate changes in odds (or chances) if the symptom is absent. Odds become smaller when LR– < 1 (between 0 and 1), lower is better to exclude.

Table 6 The influence of the presence of marked sensitivity to injustice on the probability that *Causticum* will be effective

	Prior chance (%)	*Injustice*	LR+/LR−	*Posterior chance (%)*
Case 1	10	++	5.79	40
Case 2	50	0	0.360	25

Consider two possibilities:

- Patient 1: One of the first things you learn from this patient is his sensitivity to injustice; he writes 10 letters a year to politicians about all kinds of perceived injustice. Will *Causticum* work in this patient?

Table 7 Interim assessment of LR's for six symptoms

Total =	1634	*a*	*c*	*b*	*d*	*LR+*	*95% CI*	*LR−*	*95% CI*
Diarrhoea2	*n* = 57								
	Arg-n	3	3	54	1574	15.07	6.49–34.99	0,517	
	Calc	2	25	55	1552	2.16		0,959	
	Ph-ac	3	6	54	1571	10.03	3,84–26.21	0,690	
Diarrhoea1+2	*n* = 201								
	Arg-n	4	2	197	1431	5.51	3.01–9.85	0,379	
	Calc	3	24	198	1409	0.90		1,014	
	Ph-ac	3	6	198	1427	2.74	1.08–6.96	0,759	
Fear2	*n* = 63								
	Anac	2	2	61	1569	13.36	4.86–36.70	0,519	
	Ars	2	6	61	1565	6.66	1.96–22.69	0,779	
	Calc	2	25	61	1546	1.95		0,962	
Fear1+2	*n* = 146								
	Anac	2	2	144	1486	5.66	2.10–15.27	0,548	
	Ars	2	6	144	1482	2.82		0,823	
	Calc	2	25	144	1463	0.83		1,017	
Grinding2	*n* = 91								
	Bell	2	4	89	1539	6.10	1.93–19.25	0,705	
	Carc	2	12	89	1531	2.60		0,907	
	Merc	4	11	87	1532	4.96	2.09–11.77	0,775	
	Tub	2	3	89	1540	7.32	2.46–21.83	0,635	
Grinding1+2	*n* = 209								
	Bell	3	3	206	1422	3.95	1.76–8.89	0,572	
	Carc	2	12	207	1413	1.12		0,983	
	Merc	6	9	203	1416	3.19	1.69–6.01	0,686	
	Tub	2	3	207	1422	3.15	1.07–9.28	0,687	
Herpes2	*n* = 81								
	Carc	2	12	79	1541	2.93		0,901	
	Nat-m	7	47	74	1506	2.77	1.34–5.72	0,913	
	Sep	3	30	78	1523	1.87		0,956	
	Sulph	2	27	79	1526	1.40		0,979	
Herpes1+2	*n* = 268								
	Carc	4	10	264	1356	1.75		0,853	
	Nat-m	13	41	255	1325	1.49		0,905	
	Sep	9	24	259	1342	1.69		0,868	
	Sulph	5	24	263	1342	1.05		0,990	
Injustice2	*n* = 163								
	Aur	2	5	161	1466	2.89		0,793	
	Bell	2	4	161	1467	3.37	1.08–10.55	0,740	
	Caust	8	6	155	1465	5.97	3.70–9.63	0,474	0,259-0,868
	Kali-bi	2	8	161	1463	2.02		0,888	
	Lyc	3	19	160	1452	1.37		0,959	
	Merc	2	13	161	1458	1.34		0,962	
	Nat-m	5	49	158	1422	0.93		1,008	
	Ph-ac	3	6	160	1465	3.39	1.33–8.63	0,739	
Injustice1+2	*n* = 345								
	Aur	4	3	341	1286	2.73		0,542	
	Bell	3	3	342	1286	2.38		0,633	
	Caust	10	4	335	1285	3.45	2.45–4.88	0,360	0,157-0,825
	Kali-bi	4	6	341	1283	1.90		0,759	
	Lyc	3	19	342	1270	0.64		1,096	
	Merc	7	8	338	1281	2.24		0,674	
	Nat-m	14	40	331	1249	1.24		0,937	
	Ph-ac	4	5	341	1284	2.12		0,703	
Loquac2	*n* = 107								

Table 7. (*continued*)

Total =	1634	*a*	*c*	*b*	*d*	*LR+*	*95% CI*	*LR−*	*95% CI*
	Hyos	2	2	105	1525	7.76	2.86–21.04	0,534	
	Lach	3	7	104	1520	4.68	1.79–12.29	0,748	
	Lyc	2	20	105	1507	1.40		0,972	
	Merc	2	13	105	1514	2.06		0,927	
	Sulph	1	28	106	1499	0.52		1,034	
Loquac1+2	*n* = 338								
	Hyos	3	1	335	1295	3.65	2.06–6.48	0,315	
	Lach	6	4	332	1292	2.93	1.75–4.91	0,503	
	Lyc	5	17	333	1279	1.10		0,974	
	Merc	3	12	335	1284	0.97		1,009	
	Sulph	2	27	336	1269	0.33		1,178	

If 95% confidence interval (95% CI) is not mentioned it includes 1.
Total of included patients is 1634, number of assessed prescriptions is 1049.
Diarrhoea2 means: strong diarrhoea from anticipation.
Diarrhoea1+2 means: mild or strong diarrhoea from anticipation.
Column a: the medicine worked and the symptom was present.
Column c: the medicine worked and the symptom was not present.
Column b: the medicine did not work (or was not prescribed) and the symptom was present.
Column d: the medicine did not work (or was not prescribed) and the symptom was not present.

- Patient 2: You have several reasons to suspect *Causticum* as one of two possible medicines. But you get no indications that he is sensitive to injustice.

It is harder to obtain significant negative LR's (Table 7). This is consistent with the fact that the meaning of the absence of a symptom is not very clear in homeopathic practice and not included in the repertory. There are only some vague indications in materia medica like '*Arsenicum album* becomes improbable when the patient is not chilly'. The first significant LR− we measured is 'Sensitive to injustice' and *Causticum*. We see that the absence of sensitivity to injustice does not make a striking difference despite the fact that this LR− is quite low compared to other LR−'s. So we cannot draw any consequences from the fact that a patient is not loquacious if we consider prescribing *Sulphur*.

Discussion

At this stage we cannot be certain about translating LR values into typefaces. The values mentioned above seem practical, but confidence intervals are large and we can only estimate the prior chance that a homeopathic medicine will work without any knowledge about the patient. But the meaning of typeface in the repertory has always been very uncertain and we suspected mistakes. These mistakes are now more evident.

There is no easy solution to this problem. We cannot simply strike all the entries of the most frequently used medicines from certain rubrics; some are still relevant. Of the four rubrics we showed above, the rubrics 'fear of death' and 'loquacity' are more defective than the grinding of teeth during sleep and herpes of the lips. More defective rubrics are also larger than the other two. It is obvious that common symptoms have larger rubrics than rare symptoms, but in this case there is not much difference in symptom-prevalence (fear of death = 3.9%, grinding teeth during sleep = 5.6%, herpes of lips = 5%, loquacity = 6.5%). This is an indication that the adding of medicines to repertory rubrics has been arbitrary.

Are our figures valid all over the world? This should be investigated further. We were able to find figures for the prevalence of the symptoms 'grinding teeth at night' (= bruxism) and 'herpes lips' from a Pubmed search. For herpes lips we found large variations: 3–33% in different populations.[4–7] This also reflected different threshold values. For frequent 'herpes lips' values came near to ours, 5.4% in a population of over 20,000 Israeli soldiers,[4] and 4% in the American army.[5] The prevalence of grinding teeth in our population was 5.6% on a weekly basis, and 12.8% reported grinding occurring more than twice a month. Variation in literature was slightly less, but again reflected different definitions and threshold values. A survey in Finland showed 3.7% on a weekly basis.[8] We found less clear definitions in surveys in the USA (7.9%),[9] Hong Kong (8.5%),[10] Canada (8%),[11] Nigeria (14.4%),[12] and Sardinia (27.2%).[13] Therefore, we think that the definition of the symptom is a bigger problem than variation across the world. But it is always necessary to mention the prevalence in the investigated population.

The handling of absence of expected symptoms is now within reach using Bayesian theory, but a new problem arises. What do we regard as absence of a symptom? This may be more difficult to define than the presence of a symptom, especially if the symptom is vague. However, we do use absence of expected symptoms in practice. We can make the discussion about this subject more concrete with Bayesian theory

and we must think about the way we can include LR– in our method and repertory.

Conclusion

Kent's repertory has serious flaws. Frequently used medicines are over-rated, seldom used medicines are under-rated. There is no clear system, so we have to overcome the problem by proper assessment of the prevalence of symptoms in populations responding to homeopathic medicines. Homeopathic doctors have unknowingly endorsed Bayesian philosophy for more than 200 years, but it is now necessary to do so more consciously. This will help us to improve the repertory and to get maximum benefits from our instruments, homeopathic symptoms, for instance the meaning of absence of symptoms. LR research of homeopathic symptoms is feasible. We expect to obtain six optimised repertory rubrics in 3 years time thanks to the efforts of 10–15 colleagues.

Acknowledgements

We thank the participating colleagues for their efforts and inspiring discussions. This research is supported by the Louise van Eeghen Foundation, the Dutch foundation for funding of scientific research SFWOH, the Dutch association of homeopathic physicians (VHAN), the Dutch homeopathic training institute for physicians (SHO) and the Dutch Royal association for homeopathy (KVHN). Conflicts of interests: none.

References

1 Kent JT. *Repertory of the Homeopathic Materia Medica.* New Delhi (India): World Homeopathic Links, 1945.
2 Hering C. *The Guiding Symptoms of Our Materia Medica.* New Delhi: B. Jain Publishers, 1974.
3 Stolper CF, Rutten AL, Lugten RF, Barthels RJ. Improving homeopathic prescribing by applying epidemiological techniques: the role of likelihood ratio. *Homeopathy* 2002; **91**: 230–238.
4 Katz J, Chaushu G, Peretz B. Recurrent oral ulcerations associated with recurrent herpes labialis – two distinct entities?. *Community Dent Oral Epidemiol* 2001;29:260–263.
5 Shulman JD, Carpenter WM, Lewis DL. The prevalence of recurrent herpes labialis during an army hot weather exercise. *J Public Health Dent* 1992; **52**: 198–203.
6 Axell T, Liedholm R. Occurrence of recurrent herpes labialis in an adult Swedish population. *Acta Odontol Scand* 1990; **48**: 119–123.
7 Embil JA, Stephens RG, Manuel FR. Prevalence of recurrent herpes labialis and aphthous ulcers among young adults on six continents. *Can Med Assoc J* 1975; **113**: 627–630.
8 Hublin C, Kaprio J, Partinen M, Koskenvuo M. Sleep bruxism based on self-report in a nationwide twin cohort. *J Sleep Res* 1998; **7**: 61–67.
9 Solberg WK, Woo MW, Houston JB. Prevalence of mandibular dysfunction in young adults. *J Am Dent Assoc* 1979; **98**: 25–34.
10 Ng DK, Kwok KL, Poon G, Chau KW. Habitual snoring and sleep bruxism in a paediatric outpatient population in Hong Kong. *Singapore Med J* 2002; **43**: 55.
11 Lavigne GL, Lobbezoo F, Rompre PH, Nielsen TA, Montplaisir J. Cigarette smoking as a risk factor or an exacerbating factor for restless legs syndrome and sleep bruxism. *Sleep* 1997; **20**: 290–293.
12 Oginni O, Olusile AO. The prevalence, aetiology and clinical appearance of tooth wear: the Nigerian experience. *Int Dent J* 2002; **52**: 268–272.
13 Melis M, Abou-Atme YS. Prevalence of bruxism awareness in a Sardinian population. *Cranio* 2003; **21**: 144–151.