FO-03 Judging a remedy without a laboratory: the evidence ladder for lay healers

HEALING after the FALL, research library, FO-03

Researched 10 October 2026 by Claude Opus agent for THRIVING after the FALL, Book Four. Library document, unverified.

1. Why the book needs one method

After the Fall a healer will be offered thousands of remedies by grandmothers, salvaged herbals, travellers and this book. Most will be harmless and useless, some will work, and a few will kill. There will be no laboratory and no trial register. What remains is careful reasoning, written records and comparison.

The method rests on three habits. Keep apart what the tradition says, what the record shows and what modern science infers. Know the traps that make useless remedies look good, because all of them operate in every village. Grade every claim on one ladder, print the grade, and keep learning from a register.

The chapter also answers a question the whole book depends on: when unrelated peoples use the same plant for the same illness, how much is that worth? In short, it roughly doubles the chance that a plant contains something biologically active, says much less about whether a cup of the tea helps a sick person, and can be faked by shared books and shared ancestors.

2. Three things to keep apart

What the tradition says is a report of belief: "foxglove cures the dropsy".

What the record shows is what happened to people who received it, counted, ideally against people who did not. William Withering's 1785 account of foxglove (*Digitalis purpurea*, leaf) reported 163 patients with dropsy treated between 1775 and 1785, failures included, and described overdose: vomiting, yellow vision and a pulse slowed to about 35 a minute (Withering 1785, as described by the James Lind Library and Hektoen International 2017).

What modern science infers is mechanism and controlled evidence: foxglove contains cardiac glycosides, the purified drug is digoxin, and the gap between a useful and a lethal dose is narrow (Hektoen International 2017).

The three often disagree. The Madagascar periwinkle (*Catharanthus roseus*, leaf) was drunk as a tea for diabetes in Jamaica. When Robert Noble's laboratory in Canada tested it in the 1950s, it did little to blood sugar in rats but lowered their white cell counts, which led to vinblastine and vincristine for leukaemia and other cancers (John Innes Centre, undated). The tradition pointed to real pharmacology, but for the wrong illness. Believing it would have meant treating diabetes with a cytotoxic plant.

3. The traps that make useless remedies look good

Each trap makes a remedy that does nothing appear to work in an honest healer's experience.

3.1 Natural recovery

Most illnesses get better on their own, and whatever was given last takes the credit. Henry Beecher claimed in 1955 that placebos relieved about 35 per cent of 1,082 patients in fifteen studies. Kienle and Kiene re-examined those studies and found that ten of the fifteen ignored spontaneous improvement. One, a 1933 study of the common cold, credited the placebo with recovery that colds achieve unaided (Kienle and Kiene 1997).

A Danish group diagnosed duodenal ulcers by endoscopy in 91 patients and simply watched. Twenty-nine (32 per cent) had healed in two weeks and 52 (75 per cent) by six weeks, on antacids and advice alone (reported in Moerman 2000). Any remedy given to those patients would have looked good.

Worked example. A healer gives a bitter tea to ten people with cough and fever, and eight are well within ten days. If eight in ten such coughs clear unaided in ten days, the tea has shown nothing. The question is never "did they get better?" but "did more get better, or faster, than without it?"

3.2 Regression to the mean

People come for help at their worst. Anything measured at an extreme tends to be less extreme next time, because extremes are partly chance. Galton described it in 1886 from the heights of 930 adult children and their parents (Morton and Torgerson 2003). Among women on osteoporosis treatment who lost bone in the first year, more than 80 per cent gained bone in the second with no change of treatment. The same authors warn that a new treatment tried first on the sickest patients will usually produce a gratifying response for this reason alone (Morton and Torgerson 2003). McDonald, Mazzuca and McCabe (1983) estimated that regression alone could produce apparent improvements of 2 to 37 per cent on laboratory measures (secondary only, and see Senn's correction of their formula).

Worked example. The day a woman walks two hours to the healer about her knee is the worst day of her month. Whatever is rubbed on it, next week will usually be better.

3.3 The placebo and the meaning response

Care itself changes how people feel. Moerman and Jonas (2002) proposed the term "meaning response" for the effects of the meaning of a treatment, and argued that natural recovery and regression should be removed from what is called the placebo effect. Measured examples:

- In January 1799 John Haygarth tested Elisha Perkins's patented "metallic tractors" against wooden copies painted to look like metal. Four of five chronic rheumatic patients reported relief from the fakes, and the real tractors did no better on the next day (Haygarth 1800, as described by the James Lind Library).
- Patients given a firm diagnosis and a confident promise of recovery reported getting better in 64 per cent of cases, against 39 per cent of those told the doctor was uncertain (Thomas 1987, reported in Moerman 2000).
- In a Boston trial of 262 adults with irritable bowel syndrome, adequate relief at three weeks was 28 per cent on a waiting list, 44 per cent with sham acupuncture from a brisk practitioner and 62 per cent with the same sham from a warm one (Kaptchuk et al. 2008).
- Across 117 ulcer drug trials in 32 countries, placebo healing averaged 35.3 per cent but ranged from 0 to 100 per cent: 59 per cent in six German trials, 22 per cent in five Danish and Dutch trials and 7 per cent in three Brazilian trials (Moerman 2000). Germany showed the smallest placebo response in blood pressure trials, so the effect varies by condition as well as culture.

The limits are just as clear. Hróbjartsson and Gøtzsche (2001) compared placebo with no treatment in 114 trials. For yes-or-no outcomes such as cure, placebo had no significant effect (relative risk 0.95, 95 per cent confidence interval 0.88 to 1.02). For pain the effect was about 6.5 mm on a 100 mm scale. Their 2010 Cochrane update of 234 trials found modest effects on reported pain (standardised mean difference minus 0.28) and nausea (minus 0.25), with only 16 trials at low risk of bias.

Rule for the book: care, confidence and ritual genuinely ease pain, nausea, fear and feeling ill. They do not close a wound, kill bacteria, replace lost fluid or set a bone. A remedy that works only on how a person feels is not worthless, but must never displace one that works on the body.

3.4 Selective memory and the last patient

Healers remember dramatic recoveries and forget quiet failures, and they overweight the most recent case. In the Philadelphia yellow fever epidemic of 1793 Benjamin Rush bled and purged heavily and defended it in print. William Cobbett accused him of killing more than he saved, and Rush won $5,000 for libel in December 1799 (Brooklyn Law School Library 2019). A court verdict does not make a treatment work. Reliable counts of Rush's own patients' deaths were not found [VERIFY].

The defence is old: write every case down before the outcome is known, and count them all. Withering's strength was that he reported his failures. David Werner's village manual adds that the more remedies a community has for one illness, the less likely any of them works. His example is goitre in rural Mexico, with many useless folk cures, while iodised salt prevents it (Werner et al. 2024).

3.5 Care confounded with cure

People who receive a remedy differ from those who do not.

- **Who follows advice.** In a large beta-blocker trial after heart attack, patients who took more than 80 per cent of their pills had five-year mortality of 15 per cent against 25 per cent for poor takers. The same split appeared among those taking placebo pills, 15 against 28 per cent (Horwitz et al. 1990, reported in Moerman 2000). The family that faithfully gives the tea is also the family that boils water and feeds the patient.
- **Who is treated first.** In 1835 Pierre Louis compared 77 pneumonia patients in Paris: 41 bled early (days 1 to 4) and 36 bled late (days 5 to 9). About 44 per cent of the early group died against 25 per cent of the late group, but the early group had nearly twice as many patients over 50 (Morabia 2004). Louis's counts did not prove bleeding harmful, but they removed the grounds for believing it helped.
- **Who adopts new advice.** When babies were put to sleep on their fronts, low-risk families adopted the advice first, which partly hid the harm (Gilbert et al. 2005).

3.6 Publication and vendor bias

What is written down is not a fair sample. Of 74 antidepressant trials registered with the United States regulator, 31 per cent were never published. In the published literature 94 per cent looked positive, against 51 per cent in the regulator's own analysis (Turner et al. 2008). In a 1998 review, every acupuncture trial from China, Japan, Hong Kong and Taiwan was positive, and Chinese trials of other treatments were positive 99 per cent of the time against about three-quarters for England (Vickers et al. 1998).

Herbals record cures, not failures, and the seller has an interest. Critics can be wrong too: a much-quoted 2013 study claiming that most North American herbal products were substituted was retracted in 2024 after an investigation found evidence of data fabrication (Newmaster et al. 2013). Sceptical evidence needs the same scrutiny as enthusiastic evidence.

3.7 Plausible theory never tested

From 1956 Benjamin Spock's best-selling baby book advised front sleeping so that babies would not choke on vomit. By 1970 the studies already showed about three times the risk of sudden infant death (pooled odds ratio 2.93). Acting then might have prevented over 10,000 deaths in the United Kingdom and at least 50,000 in Europe, the USA and Australasia, a counterfactual estimate rather than a count (Gilbert et al. 2005). A sensible mechanism is a reason to test, not a reason to believe.

4. The evidence ladder

Grade	Name	What it requires	How to act
A	Tested fairly	Randomised trials or a review of them with a clear result	Use if you have the means and the patient fits
B	Counted and consistent	Consistent clinical series, or a dramatic effect seen repeatedly	Use, and record
C	Convergent and plausible	Same use for the same purpose by unrelated cultures, plus active chemistry that fits	Worth a careful, recorded trial if safe. Not proof
D	Claimed	One tradition, one healer, anecdote, or theory alone	Do not rely on it. Test only if harmless
X	Failed or harmful	Fails fair tests, or harms	Do not use

Rule 1. Grade the claim, not the plant. Periwinkle tea for diabetes is D at best, while the alkaloids from the same leaf against leukaemia rest on trials.

Rule 2. Harm overrides benefit. Credible serious harm at ordinary doses means X for that use. When a Belgian slimming clinic's *Stephania tetrandra* root was replaced by *Aristolochia fangchi* root, over 100 people developed kidney failure, and 18 of 39 who had kidneys and ureters removed (46 per cent) had urothelial cancer. Cumulative doses above 200 g carried higher risk (Nortier et al. 2000). Bradford Hill made the same point: slight evidence may justify stopping something that may harm, while strong evidence is needed before asking people to do something burdensome (Hill 1965).

Rule 3. A dramatic effect can earn B without a trial. Glasziou, Chalmers, Rawlins and McCulloch (2007) argued that when an effect is large and fast compared with the expected course, bias and chance become implausible. They proposed a rate ratio of at least 10 with a probability below 0.01, noting that in their examples it was at least 20, and listed insulin, blood transfusion and defibrillation. GRADE allows observational evidence to be rated up one level for a two-fold change in risk and two levels for five-fold (Guyatt et al. 2011). Semmelweis is the village-scale model: maternal deaths on the doctors' ward in Vienna were 18.3 per cent in April 1847, and after chlorinated lime handwashing began in May, 2.2 per cent in June and 1.2 per cent in July (James Lind Library, Semmelweis entry).

Rule 4. Modest effects need fair comparison. When the relative risk is about 2 or less, records mislead. Hormone replacement therapy and beta carotene looked protective in observational studies and failed in trials (Glasziou et al. 2007). For effects this size the book says "probably", not "does".

Rule 5. Print three columns and a date. Each graded entry gives what the tradition says, what the record shows and what science infers, with the date of the last review. A register may move a grade up or down.

The ladder ranks how far a claim has been tested, not cultures. A C-grade plant that grows in the hedge may serve better than an A-grade drug the village cannot obtain. And the published A-grade record can itself be inflated, as Turner's data show.

5. Testing the ladder: how much is convergent use worth?

Grade C rests on the idea that independent discovery is evidence. This section asks how far that holds, in numbers.

5.1 Fabricant and Farnsworth: the 80 per cent figure

Fabricant and Farnsworth (2001) identified 122 compounds of defined structure, from only 94 species, used as drugs somewhere in the world, and stated that 80 per cent had an ethnomedical use identical or related to the drug's current use. An earlier version reported 119 compounds and, according to secondary sources, 74 per cent (Farnsworth et al. 1985) [VERIFY]. Only the abstract of the 2001 paper could be read, so the meaning of "related" could not be checked.

The figure is often misread. It says that among plants that became drugs, most had a matching traditional use. It does not say that most traditional remedies work. This is the hit-rate denominator problem. Kew recorded at least 28,187 plant species with documented medicinal use (Royal Botanic Gardens Kew 2017). If all 94 drug species are among them, about one recorded medicinal species in 300 (0.3 per cent) has yielded a globally used pure compound [author's arithmetic, VERIFY]. That is a harsh test, since a crude remedy can help without becoming a drug, but it shows the size of the denominator. The often-quoted claim that only about 6 per cent of plants have been screened for activity could not be traced to its source [VERIFY].

5.2 Moerman: medicinal floras are not random

Daniel Moerman asked whether peoples choose medicinal plants at random from what grows around them, by regressing medicinal species per plant family on total species per family. In North America north of Mexico, 2,147 of 16,270 species were used medicinally by 123 societies in 17,624 recorded ways (Moerman 1991, abstract). In the 1999 comparison the regression predicted that 306.7 species of the daisy family (Asteraceae) would be used if choice were random, but 397 were. For grasses (Poaceae), 171.8 were predicted and 62 used (Moerman et al. 1999).

Comparing North America with Kashmir, Korea and the Chiapas highlands, the family-by-family residuals correlated strongly (0.607 to 0.743). Asteraceae ranked first in three of four regions and second in the fourth, and grasses last in three and second-last in the fourth. Eastern Ecuador correlated weakly with the others (minus 0.043 to 0.189). Moerman offered two explanations: related floras share related chemistry, and related peoples may have carried knowledge across Asia into the Americas (Moerman et al. 1999). Yarrow (*Achillea millefolium*, flowering tops), with 359 recorded North American uses, is the most used species.

Two warnings. Bennett and Husby (2008) showed the regression method is statistically flawed, and with better methods only 10 of 115 families in the Shuar flora of Ecuador departed significantly from chance. And families rich in defensive chemistry will be used for everything, which signals activity but not the right activity.

5.3 Cox in Samoa

Paul Alan Cox worked with Samoan healers from the 1980s. Extracts of 74 healer-nominated species were tested in a broad animal screen and a guinea-pig gut test, and over 86 per cent showed activity (Cox et al. 1989, abstract, authors [VERIFY]). One lead, the tree *Homalanthus nutans* (mamala, stem wood and bark), used by a healer for a hepatitis-like illness, yielded prostratin, studied against HIV. In 2004 Samoa and the University of California, Berkeley agreed to share proceeds from a gene sequence for producing it (UC Berkeley 2004). Cox and Balick (1994) argued that ethnobotanical selection beats random collection.

The Samoan figure has no control group, and a broad screen detects almost any biological effect. It shows that healer-chosen plants are active, not that they beat chance, nor that they help the named illness.

5.4 Head-to-head comparisons

The fair test screens healer-chosen and random plants in the same laboratory. Albuquerque and colleagues (2014) summarised the main attempts:

- Sinai: 83 per cent of traditionally used plants were antimicrobial against 42 per cent of random plants, but more random plants were strongly active (13.9 against 8.3 per cent).
- Belize: 4 of 31 traditionally used plants relaxed blood vessel muscle, against none of 32 random plants.
- Medicinal and poisonous plants yielded anticancer activity at 1.4 to 2.6 times the rate of others.

Gyllenhaal and colleagues (2012) compared plants with recorded uses in Vietnam and Laos with plants collected for diversity in a Vietnamese national park. Per collection, random plants had the higher hit rate. Per sample, traditional plants did better, notably against tuberculosis and malaria, with little advantage for HIV or cancer. Plants whose traditional use matched the specific assay did no better than random ones (abstract).

5.5 Phylogeny as an independence check

Saslis-Lagoudakis and colleagues (2012) built a family tree of about 20,000 species from Nepal, New Zealand and the Cape of South Africa, of which 1,470 were medicinal. Medicinal plants clustered: "hot" branches held 60 per cent more of them than random samples. Across the three regions, which share almost no recent cultural contact, hot branches from one region held 17 per cent more medicinal plants in the others than expected, and 38 per cent more for matched conditions. They also held more genera that have yielded drugs. The authors note that bioactivity was inferred from drug genera, not tested. This is the best current evidence that cross-cultural convergence reflects something real in the plants.

5.6 Shared books are not independent discovery

Marco Leonti (2011) argued that texts spread plant knowledge further and longer than speech, so field surveys may re-record what came from a book. Later work traced present-day use in southern Italy partly to Dioscorides and Galen (Leonti et al. 2015). After the Fall the same will happen with this book. Within a generation its advice will appear as "tradition" across a district, and the book must never count its own echoes as confirmation.

5.7 The cost of a hit within a good tradition

Tu Youyou's antimalarial search began in January 1969 with more than 2,000 candidate remedies from Chinese texts and practitioners, summarised as 640 prescriptions. By 1971 her group had made about 380 extracts from about 200 herbs. Sweet wormwood (*Artemisia annua*, leaf) gave inconsistent results below 50 per cent inhibition. Re-reading Ge Hong's fourth-century emergency handbook, which says to soak the herb and wring out the juice rather than boil it, she switched to cold ether extraction. Sample 191, made after 4 October 1971, gave 100 per cent inhibition in rodents, and human testing followed in 1972 (Asimov Press, undated). Sources differ on the volume in Ge Hong's recipe and on where the first 21 patients were treated [VERIFY].

This is tradition vindicated. It is also a measurement: hundreds of preparations, one that changed medicine, and that one only after the method was corrected.

5.8 The verdict, in numbers

Convergent or healer-nominated use roughly doubles the chance that a plant shows activity in a laboratory screen, with ratios of about 1.4 to 2 in the clearest studies and some reversals [author's synthesis of Albuquerque et al. 2014, Gyllenhaal et al. 2012 and Saslis-Lagoudakis et al. 2012, VERIFY]. It is weakest at predicting activity against the specific illness named. Activity in a dish is far from benefit at the bedside. So grade C means "more likely than a random plant to contain something active, worth a careful trial if safe". It never means "works".

6. Bradford Hill in plain words

In 1965 Austin Bradford Hill set out nine aspects to weigh before calling an association cause and effect. He called them viewpoints and said none is indispensable (Hill 1965). For a healer judging a remedy:

1. **Strength.** Is the difference big? Lung cancer deaths were nine to ten times higher in smokers and twenty to thirty times in heavy smokers. Deaths falling from 18 to 2 in 100 is strong. A cough lasting six days instead of seven is not.
2. **Consistency.** Seen by different people, in different places? Five villages' registers agreeing beats one healer.
3. **Specificity.** One condition, or "everything"? A cure-all is suspect, though one cause can have many effects.
4. **Temporality.** Did improvement follow the remedy? If the fever was already falling, the remedy gets no credit.
5. **Dose and response.** Does more, within safe limits, give more effect? Nortier's 200 g threshold for *Aristolochia* is this viewpoint applied to harm.
6. **Plausibility.** Is there a sensible way it could work? Useful but limited. Spock was plausible and wrong.
7. **Coherence.** Does it fit the known course of the illness? A "cure" for a self-limiting cold fits natural recovery, which is the problem.
8. **Experiment.** Start or stop it deliberately, and does the outcome follow? For a healer, the N-of-1 trial and planned comparison are experiments.
9. **Analogy.** Does a similar remedy work for a similar illness? Weak alone, useful for choosing what to test.

Hill added that significance tests cannot settle these questions, that the evidence needed should depend on what is at stake, and that "All scientific work is incomplete", which does not license ignoring what we know.

7. Tools a lay healer can use to learn over years

7.1 The before-and-after record

Measure, give the remedy, measure again. It is the weakest design, because natural recovery, regression and meaning all produce improvement. It becomes useful when the condition has been stable and not improving, the measure is objective (a traced wound, a counted pulse, a weight), and the change is large and fast. Semmelweis's monthly death rates pass all three tests.

7.2 The N-of-1 trial

For one patient with a chronic, stable condition, alternate. Guyatt and colleagues at McMaster University introduced this in 1986: pairs of periods, remedy and dummy, order within each pair by chance, patient unaware which is which. Their first patient, an asthmatic, was worse on theophylline (Guyatt et al. 1986, secondary). In three years 57 such trials were completed, 50 gave a definite answer, and 15 changed the plan, 11 by stopping a drug (Guyatt et al. 1990). The 39 per cent change rate often quoted refers only to trials with usable data, and is about 26 per cent of completed trials [author's arithmetic].

It needs one patient, a notebook, two look-alike jars and a coin. It works only for conditions that return when treatment stops and do not threaten life: chronic pain, sleep, indigestion, itch, chronic cough, low mood. Never use it for an acute infection.

7.3 The village register

One line per case, written at first contact: date, number, age, sex, condition by the book's case definition, severity on a fixed scale, remedy and amount, then the outcome on a fixed day. Louis's numerical method was no more than this, tabulated and compared, though as Morabia (2004) notes, groups only become comparable when allocation is by chance.

The register defeats selective memory, provides a baseline and, pooled, provides numbers. Detecting a fall in deaths from 40 to 20 per cent with conventional confidence needs about 80 patients per group [author's calculation, VERIFY]. Few conditions produce that in a village of 300. A district of 3,000 pooling ten registers may manage it for common illnesses within a year or two.

7.4 The planned comparison

When the healer honestly does not know which of two reasonable remedies is better, give them by lot to successive comparable patients and record outcomes. On HMS Salisbury from 20 May 1747, James Lind gave twelve sailors with scurvy six different additions to the same diet, two men each. The pair given two oranges and one lemon daily recovered fastest, one fit for duty within six days (Lind 1753, summarised by Katch, undated). Lind chose the pairs himself, and two per arm is tiny, but the design is within any village's reach.

A comparison is justified only when both options are acceptable, the patient or family agrees, and it stops if one arm does clearly worse. Never withhold something known to work, such as oral rehydration for a child with diarrhoea.

8. Reading a herbal, a manual or a village elder

8.1 Seven questions for any claim

1. **What exactly is claimed?** Illness, plant, part, preparation, amount, frequency, duration, and expected result by when.
2. **What is the plant, precisely?** Local name and a pressed specimen. The *Aristolochia* disaster began with two plants sharing one trade name.
3. **How do you know?** Seen, heard or read, and how often?
4. **Who failed or was harmed?** A healer who reports no failures has not been counting.
5. **Who else uses it?** Unrelated peoples, or only this lineage and its books?
6. **Who gains?** Is it sold, secret, or tied to the healer's standing?
7. **What would make you stop?** A claim nothing could disprove is not knowledge.

8.2 Werner's four rules

Where There Is No Doctor gives four rules from Mexican villages (Werner et al. 2024, chapter 1): the more remedies for one illness, the less likely any works. Foul or disgusting remedies seldom help and often harm. Remedies of animal or human dung can cause dangerous infections, including tetanus on a newborn's navel, and must never be used. The more a remedy resembles the illness, the more likely its benefit comes from belief alone.

These rules are themselves untested, grade D. They protect against real harm, but will sometimes be wrong. Rule one and the convergence principle point opposite ways: many remedies within one culture suggests none works, one remedy across many cultures suggests it might.

8.3 Counting agreement among healers

The informant consensus factor (Trotter and Logan 1986, adapted by Heinrich et al. 1998) is the number of use reports for an illness category, minus the number of different plants named, divided by the number of use reports minus one. It runs from 0 (everyone names a different plant) to 1 (everyone names the same). Heinrich's Mexican study found high agreement for gut complaints. High agreement within one culture is evidence of shared belief, not of effect, but it tells a district which remedies to test first.

Evidence grades

Grades for the methods and the historical examples in this chapter.

Item	Grade	Tradition says	Record shows	Science infers
Healer's unrecorded impression	D	Experience teaches	Distorted by recovery, regression, memory (Morton and Torgerson 2003)	Unreliable for modest effects
Placebo changes cure or death	X	Belief heals	No effect on binary outcomes (Hróbjartsson and Gøtzsche 2001)	Effects modest, on reported symptoms
Meaning response eases pain and nausea	A	Care heals	SMD about minus 0.25 (Hróbjartsson and Gøtzsche 2010)	Real, modest, variable
Warm, confident care improves reported symptoms	A for IBS, B generally	Manner matters	28, 44, 62 per cent (Kaptchuk et al. 2008)	Real for subjective outcomes
Convergent use predicts laboratory activity	B	Many peoples cannot be wrong	Roughly double, inconsistent (Albuquerque et al. 2014, Gyllenhaal et al. 2012)	Partly phylogenetic (Saslis-Lagoudakis et al. 2012)
Convergent use predicts clinical benefit	C	As above	No direct tests found [VERIFY]	Plausible, unmeasured
N-of-1 trial, chronic stable conditions	B	Not traditional	50 of 57 definite answers (Guyatt et al. 1990)	Sound where symptoms return off treatment
Before-and-after, fluctuating condition	D	Common practice	Misleads (Kienle and Kiene 1997)	Invalid without baseline
Before-and-after, dramatic effect on stable baseline	B	None	Semmelweis, 18.3 to 1.2 per cent	Valid near rate ratio 10 (Glasziou et al. 2007)
Werner's four rules	D	Village experience	Not tested	Consistent with infection risk and placebo work
Citrus for scurvy	B	Sailors' lore	Lind 1747	Vitamin C deficiency
Foxglove for dropsy	B, narrow margin	Shropshire folk remedy	163 cases (Withering 1785)	Cardiac glycosides
Handwashing with chlorinated lime at births	B	None	Tenfold fall (Semmelweis)	Germ transmission
Early bloodletting for pneumonia	X	Standard to 1830s	44 against 25 per cent deaths, confounded (Morabia 2004)	No benefit

Item	Grade	Tradition says	Record shows	Science infers
Metallic tractors	X	Sold as cure	Wooden fakes equal (Haygarth 1800)	No mechanism
Aristolochia root, internal use	X	Several traditions	18 of 39 cancers (Nortier et al. 2000)	Nephrotoxic, carcinogenic
Front sleeping for infants	X	Plausible theory	OR 2.93 by 1970 (Gilbert et al. 2005)	Harmful
Periwinkle tea for diabetes	D	Jamaican use	Little effect on rat blood sugar	Cytotoxic alkaloids

Means ledger

Rung	What can be done	What it changes
Hands and knowledge	Know the traps, ask the seven questions, apply Werner's rules and red flags, count pulse and breaths	Stops the worst remedies and delays. Cannot detect modest effects
Hearth	A family notebook, before-and-after records (a string round a swelling, a wound traced on cloth), N-of-1 trials with two look-alike jars and a coin	One household learns which remedies help its chronic sufferers
Garden and still	Weighed doses, so dose and response can be seen. Look-alike dummy preparations. Pressed voucher specimens of every plant used	Reproducible doses. Catches substitution. Makes Hill's fifth viewpoint visible
Village dispensary	Bound register, fixed case definitions and outcome days, thermometer, child scale, microscope where salvaged so remedies are judged on the right illness	Removes false credit from look-alike illnesses, gives baselines. Too few cases for modest effects
District infirmary	Pooled registers, allocation by lot with consent, blinded assessor, annual grade review	About 80 per arm becomes reachable, so a halving of deaths in a common illness can be seen in a year or two
Hospital before the Fall (reference)	Randomised trials, reviews, registries, chemical and DNA identity testing	Detects effects of a few per cent. Still open to publication bias (Turner et al. 2008)

Harm test

False confidence. Natural recovery and meaning make most remedies look good, and the harm is delay of something that works or of sending the patient on. Every remedy entry should list the conditions it must never delay.

False rejection. Ten patients cannot show a modest benefit. Never downgrade an A or B remedy on a village's small series.

Testing toxic plants. Foxglove, *Aristolochia* and periwinkle show that "active" often means "dangerous". This chapter gives no method for testing an unknown plant in a person. N-of-1 trials and comparisons use only remedies

graded C or better and listed safe at the dose. Never test on pregnant women, infants, the very old, the dehydrated, anyone with kidney or liver disease, or anyone acutely ill.

Misidentification. Trade names, common names and powders hide the plant. Keep voucher specimens and use whole, recognisable material.

A contaminated record. A register written after the outcome is selective memory on paper. Write at first contact.

Withholding care in a comparison. Compare only under genuine uncertainty, never withhold A or B care, and stop when one arm does clearly worse.

Breach of confidence. Pooled copies carry numbers, not names.

Loss of comfort. Where a D remedy is harmless and the illness self-limiting, the healer may let it be, provided nothing better is withheld.

Historical harms from bad judgement include Rush's purging in 1793, bloodletting for pneumonia, front-sleeping advice from 1943 to 1988, and the Belgian slimming clinic of the early 1990s.

Red flags

The commonest misuse of this chapter would be to "try a remedy and see" in someone who has no time. Do not test anything when:

- the person cannot keep fluids down, has stopped passing urine, or is drowsy or confused
- breathing is fast or laboured at rest, or the lips are blue
- there is a stiff neck with fever, or a bulging soft spot in a baby (Werner et al. 2024)
- bleeding will not stop, or a pregnant woman bleeds
- a wound has spreading redness, a foul smell or crackling under the skin
- a child is floppy, will not feed or cannot be woken properly
- the illness worsens day by day despite care.

What to do anyway: give what the book grades A or B, keep the person warm, hydrated if they can swallow, clean and accompanied, and send for the most skilled help within reach. Where nothing more can be done, honest comfort is the right care: relief of pain and thirst, company, clear words and no false promise. The meaning response is real, and in the last days it is the treatment. The specific clinical red flags belong to the clinical chapters and should be cross-checked there.

Red flags in a claim: a cure for everything, a seller, a secret, no failures, testimonials only, "no side effects", forbids other treatment, discourages records.

The case against

Against grading. Letters invite false precision. Jürg Gertsch (2009) argued that ethnopharmacology over-reads laboratory data and that reproducible does not mean meaningful. Two healers may grade the same evidence differently. *Answer:* print the three columns behind each grade so readers can disagree.

Against the ethnobotanical signal. Gyllenhaal and colleagues (2012) found random collection sometimes did better, and matching use to assay gave no advantage. Bennett and Husby (2008) showed Moerman's regression is flawed. Leonti (2011) showed convergence can be manufactured by texts. The 80 per cent figure looks backwards. *Answer:* convergence is a reason to test, capped at grade C.

Against the trial hierarchy. Ted Kaptchuk (1998) argued that the randomised trial strips out the meaning and relationship that may be part of healing, and that the "powerful placebo" was partly a product of the method. His own trial moved reported relief from 44 to 62 per cent by manner alone. And trials are distorted by selective publication

(Turner et al. 2008, Vickers et al. 1998). *Answer:* the book grades the care itself as a real intervention for symptoms, and reads the published record with the same suspicion as the herbal.

Against the placebo sceptics. Howard Fields and others criticised Hróbjartsson and Gøtzsche for pooling unlike trials, and Senn corrected McDonald's regression formula. The size of the meaning response is still disputed. *Answer:* claim only modest effects on reported symptoms and none on hard outcomes.

Against lay testing. Guyatt's group began N-of-1 trials because individual clinical impression is so biased. Village numbers are small and registers will mostly yield noise. *Answer:* pool at district level, rely on the dramatic-effect rule, and never downgrade A or B on village data alone.

Against romance on either side. Old is not safe: *Aristolochia* has centuries of use. Modern is not right: front sleeping had the authority of the world's best-selling baby book. Rejecting tradition would have missed artemisinin. Accepting it wholesale would have kept the bleeding bowl.

Today and after the Fall. Today a reader can consult a Cochrane review in minutes. After the Fall the only evidence will be this book and what the registers accumulate, so the method leans on dramatic effects, records and pooling. Subtle effects will not be detectable for a generation.

Plants named in this chapter

Identification is safety-critical. Lookalikes marked [VERIFY] were not confirmed from two sources here and must be checked against the plant monographs.

Plant	Part	Why named	Lookalikes or substitutes
Digitalis purpurea (foxglove)	Leaf	Withering	Comfrey (*Symphytum officinale*) leaf before flowering [VERIFY]
Catharanthus roseus (Madagascar periwinkle)	Leaf	Wrong-target tradition	Not assessed [VERIFY]
Aristolochia fangchi	Root	The dangerous substitute	Sold under a trade name shared with *Stephania tetrandra* (Nortier et al. 2000)
Stephania tetrandra	Root	Intended slimming herb	*Aristolochia* species (Nortier et al. 2000)
Artemisia annua (sweet wormwood)	Leaf	Artemisinin	Other *Artemisia* species with little or no artemisinin [VERIFY]
Homalanthus nutans (mamala)	Stem wood, bark	Prostratin	Not assessed [VERIFY]
Achillea millefolium (yarrow)	Flowering tops	Moerman	White-flowered carrot family, including poison hemlock (*Conium maculatum*) [VERIFY]

10-line summary

1. Most remedies look as if they work because illness improves unaided, people seek help at their worst, and care eases symptoms.

2. Placebo and care have small real effects on pain and nausea (SMD about minus 0.25) and none detected on cure or death.
3. Keep apart what the tradition says, what the record shows and what science infers, and print all three.
4. Grade claims A, B, C, D or X. Harm overrides benefit. A dramatic effect (rate ratio about 10) can earn B without a trial.
5. Convergent traditional use roughly doubles the chance of laboratory activity, with exceptions, and says little about the named illness.
6. The 80 per cent figure looks backwards from drugs. About 0.3 per cent of recorded medicinal species have yielded a global drug compound [inference].
7. Shared books and ancestry can fake convergence. The book must not count its own echoes.
8. N-of-1 trials gave a definite answer in 50 of 57 completed cases and changed treatment in 15.
9. A register written at first contact and pooled across a district is the only way lay healers can learn which remedies work.
10. Published science is biased too: 94 per cent of published antidepressant trials looked positive against 51 per cent in the regulator's data.

Open questions

- How often does a convergent traditional remedy help the named illness in people, as opposed to a laboratory screen? No estimate found [VERIFY].
- How large is the meaning response where biomedicine is absent, and with traditional healers?
- How should the book handle remedies where the plant is right and the preparation wrong, as with boiled qinghao?
- How quickly will this book's advice become "tradition" and contaminate the convergence signal?
- Can a district run comparisons that are ethical and accepted, and what literacy does a reliable register need?
- Werner's four rules have not been tested.
- The source of the "6 per cent screened" statistic remains untraced.

Sources to read in full

1. Evans, Thornton, Chalmers and Glasziou (2011), *Testing Treatments*, free online. Fair tests for lay readers.
2. Hill (1965). Short, and still the best guide to judging causes without experiments.
3. Glasziou et al. (2007). The rule for dramatic effects.
4. Moerman (2002), *Meaning, Medicine and the "Placebo Effect"* (not read here), with Moerman (2000).
5. Moerman et al. (1999) and Saslis-Lagoudakis et al. (2012), on convergence.
6. Gyllenhaal et al. (2012), the largest head-to-head screen.
7. Werner et al. (2024), chapter 1.
8. Leonti (2011), on texts and tradition.

At the bedside after the Fall

Household. Know the traps, never use dung or foul remedies, never give an unknown plant by mouth, keep a dated notebook for each chronic sufferer, and run an N-of-1 trial before committing to a remedy for months. *Needs:* something to write with, two identical jars, a coin, string or cloth for tracing, this chapter. *Skills:* counting pulse and breaths, knowing the red flags. *First failure:* the notebook lapses after a week. Keep each entry to one minute.

Village of 300. One bound register, fixed case definitions for the ten commonest complaints, outcome on a fixed day (for example day 7 for acute illness and day 28 for wounds and chronic conditions [VERIFY against clinical chapters]), grades beside every remedy, voucher specimens, a yearly count. *Needs:* paper, ink, a plant press, thermometer, child

scale, microscope if salvaged. *Skills:* one trained register keeper. *First failure:* outcomes go unrecorded for patients who do not return, who are disproportionately the dead and the recovered. Send someone on the outcome day.

District of 3,000. Pool registers quarterly, review grades yearly, run comparisons by lot only under genuine uncertainty and with consent, use a blinded assessor, and send a written summary back to every village. *Needs:* a clerk, a courier route, a meeting of healers. *Skills:* tallying rates, the dramatic-effect rule, the ethics of comparison. *First failure:* villages define conditions differently, so pooled numbers compare unlike with unlike. Fix and print case definitions first.

Protocols

Protocol 1. The one-page test (for printing)

JUDGING A REMEDY

A. Danger? If any red flag is present, give only A or B care and seek help. Test nothing.

B. Name the claim. - [ ] Illness: - [ ] Plant or substance, Latin name if known, part: - [ ] Preparation, amount, how often, how long: - [ ] Expected result, by when:

C. Harm. Any "yes" means do not use. - [ ] Made from dung, urine, rotting or foul matter? - [ ] Graded X in this book, or known to harm kidney, liver or heart? - [ ] Plant identity uncertain, or a powder of unknown origin? - [ ] Patient pregnant, infant, very old, dehydrated or acutely ill, and remedy not listed safe for them? - [ ] Will it delay something graded A or B?

D. Traps. Would they have got better anyway? - [ ] Does this illness clear on its own? In how many days? - [ ] Was it started at the worst moment? - [ ] Could care or ritual explain the change? - [ ] Are we remembering only successes? - [ ] Did those who got it differ from those who did not? - [ ] Does the teller gain from it?

E. Grade. A: fair trials show it works. B: counted cases show it consistently, or the effect is dramatic (about ten times, fast, on a stable baseline). C: unrelated peoples use it for the same illness and it has fitting chemistry. D: one tradition, healer, story or theory. X: failed fair tests or harms.

F. Decide. A or B: use at the book's dose. C: only if harmless at the stated dose and nothing better exists, recording every case. D: do not rely on it. X: do not use.

G. Record date, patient number, illness, severity, remedy and amount, and outcome on the fixed day.

Protocol 2. Before-and-after record

- ☐ Only for a condition stable, not improving, for at least four weeks [VERIFY, author's suggestion].
- ☐ One countable measure: wound traced on cloth, nights woken, resting pulse counted for 60 seconds, weight.
- ☐ Measure on three separate days before starting and use the middle value, which reduces regression to the mean (Morton and Torgerson 2003).
- ☐ Give the remedy at the monograph's fixed dose and period.
- ☐ Measure three times again and use the middle value.
- ☐ Count it a success only if the change is large compared with the variation seen before.
- ☐ Record failures as carefully as successes.

Protocol 3. N-of-1 trial

Only for chronic, stable, non-dangerous symptoms that return off treatment, with a remedy graded C or better and listed safe at the dose.

- ☐ Agree the daily measure, for example pain 0 to 10 each evening.

- ☐ Prepare remedy and dummy that look, smell and taste alike, labelled only "1" and "2", with the key kept by someone else.
- ☐ Period length as the monograph states, typically one to two weeks [VERIFY against monographs], with a gap between periods if the remedy lingers.
- ☐ At least three pairs [VERIFY against McMaster practice]. Toss a coin for order within each pair.
- ☐ Patient records daily without knowing which is which.
- ☐ Open the key at the end. Continue if the remedy wins all three pairs, stop if it wins none or one.
- ☐ Stop at once for serious worsening or listed side effects.
- ☐ Enter the result in the register.

Protocol 4. Village register

- ☐ Bound book, numbered pages, nothing torn out, permanent ink or pencil.
- ☐ Columns: entry, date, patient number (names kept apart), age, sex, village, condition by case definition, severity, remedy, amount and frequency, outcome date, outcome (well, better, same, worse, referred, died).
- ☐ Write everything but the outcome at first contact.
- ☐ Print the outcome day for each condition at the front.
- ☐ Visit anyone who does not return on the outcome day.
- ☐ Quarterly: count cases and outcomes per condition and remedy, copy without names for the district.
- ☐ Yearly: compare with last year and look for dramatic changes.

Protocol 5. Interviewing an elder or reading a herbal

- ☐ Write the claim in the exact words used, then complete Protocol 1, step B.
- ☐ Press a fresh specimen, labelled with date, place, local name and informant.
- ☐ Ask how often it worked, whether it ever failed or harmed, and who taught it, and whether they read it somewhere.
- ☐ For a book: author, date, sources, and whether it gives failures and doses.
- ☐ Ask at least five other healers separately, from unrelated peoples if possible, and count how many name the same plant for the same illness.
- ☐ Never count a claim from this book as independent confirmation.

Protocol 6. Planned comparison in the district

- ☐ Only under genuine uncertainty between two acceptable remedies.
- ☐ Fix case definition, outcome and outcome day before starting.
- ☐ Explain and obtain agreement from each patient or family.
- ☐ Allocate by drawing from a bag prepared in advance with equal numbers of two marks, kept by someone who does not treat.
- ☐ The person judging outcome should not know the allocation.
- ☐ Review every 20 patients and stop if one arm has clearly more deaths or serious harm.
- ☐ Aim for about 80 per arm for an important outcome [author's calculation, VERIFY]. Smaller comparisons only detect dramatic differences.
- ☐ Report the result to every village whatever it shows.

Numbers to carry

1. **32 and 75 per cent.** Ulcers healed at two and six weeks with no active treatment (Moerman 2000). A remedy must beat natural recovery, not merely precede it.

2. **Over 80 per cent** of patients who seemed to fail improved at the next measurement unchanged (Morton and Torgerson 2003). Measure three times.
3. **None detectable** effect of placebo on cure or death. About a quarter of a standard deviation on pain and nausea (Hróbjartsson and Gøtzsche 2001, 2010).
4. **Rate ratio 10.** An effect of ten times or more, fast, on a stable baseline, is very likely real (Glasziou et al. 2007). Below about 2, only a fair comparison can tell.
5. **About double.** Convergent use roughly doubles the chance of laboratory activity [synthesis, VERIFY].
6. **0.3 per cent.** About 1 in 300 recorded medicinal species has yielded a global drug compound [inference, VERIFY].
7. **80 per group** to see deaths halved from 40 to 20 per cent [author's calculation, VERIFY]. A district, not a village.
8. **50 of 57.** N-of-1 trials that gave a clear answer (Guyatt et al. 1990).
9. **94 against 51 per cent.** Published against full trial record (Turner et al. 2008). Ask who failed.

References

1. Albuquerque UP, Medeiros PM, Ramos MA, Ferreira Júnior WS, Nascimento ALB, Torres-Avilez WM, Melo JG (2014). Are ethnopharmacological surveys useful for the discovery and development of drugs from medicinal plants? Revista Brasileira de Farmacognosia 24(2). https://www.scielo.br/j/rbfar/a/TXm8c3gBn6jtnHtc5wnQB9h/?lang=en (full text read. Source for Khafagi and Dewedar 2000, Slish et al. 1999 and Spjut 2005, which were not read directly.)
2. Asimov Press (undated). Discovering an antimalarial drug in Mao's China. https://press.asimov.com/articles/antimalarial-drug (read in full. Author not recorded.)
3. Bennett BC, Husby CE (2008). Patterns of medicinal plant use: an examination of the Ecuadorian Shuar medicinal flora using contingency table and binomial analyses. Journal of Ethnopharmacology 116(3), 422-430. https://discovery.fiu.edu/display/pub79785 (abstract only.)
4. Brooklyn Law School Library (2019). Law, medicine and early American libel. https://guides.brooklaw.edu/blogs/LibraryOverview/bls-library-blog/law-medicine-and-early-american-libel (secondary summary, read via search extract.)
5. Cox PA et al. (1989). Pharmacological activity of the Samoan ethnopharmacopoeia. Economic Botany. https://link.springer.com/doi/10.1007/BF02935923 (abstract only via search extract. Author list, volume and pages [VERIFY].)
6. Cox PA, Balick MJ (1994). The ethnobotanical approach to drug discovery. Scientific American 270(6), 60-65. https://rauropiwhakaoranga.landcareresearch.co.nz/references/e99a1e16-bd02-46c0-9626-d4c8b5e7abd2 (bibliographic record only.)
7. Fabricant DS, Farnsworth NR (2001). The value of plants used in traditional medicine for drug discovery. Environmental Health Perspectives 109 (Supplement 1), 69-75. https://pmc.ncbi.nlm.nih.gov/articles/PMC1240543 (abstract only. Full text is a scanned PDF that could not be read.)
8. Farnsworth NR, Akerele O, Bingel AS, Soejarto DD, Guo Z (1985). Medicinal plants in therapy. Bulletin of the World Health Organization 63(6), 965-981. https://pmc.ncbi.nlm.nih.gov/articles/PMC2536466/ (not read. The 119 compounds and 74 per cent figures come from secondary citations.)
9. Gertsch J (2009). How scientific is the science in ethnopharmacology? Historical perspectives and epistemological problems. Journal of Ethnopharmacology 122(2), 177-183. https://boris.unibe.ch/31623 (abstract only.)
10. Gilbert R, Salanti G, Harden M, See S (2005). Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. International Journal of Epidemiology [VERIFY journal details]. https://pubmed.ncbi.nlm.nih.gov/15843394/ (abstract and secondary summaries.)

11. Glasziou P, Chalmers I, Rawlins M, McCulloch P (2007). When are randomised trials unnecessary? Picking signal from noise. BMJ 334, 349-351. https://pmc.ncbi.nlm.nih.gov/articles/PMC1800999 (full text read.)
12. Guyatt G, Sackett D, Taylor DW, Chong J, Roberts R, Pugsley S (1986). Determining optimal therapy: randomized trials in individual patients. New England Journal of Medicine 314(14), 889-892. https://www.jameslindlibrary.org/articles/history-development-n-1-trials/ (secondary accounts only.)
13. Guyatt GH, Keller JL, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT (1990). The n-of-1 randomized controlled trial: clinical usefulness. Our three-year experience. Annals of Internal Medicine [VERIFY journal]. https://pedpalascnetlibrary.omeka.net/items/show/12406 (abstract via secondary sources.)
14. Guyatt GH et al. (2011). GRADE guidelines 9: rating up the quality of evidence. Journal of Clinical Epidemiology 64(12), 1311-1316. https://edoc.unibas.ch/30223/ (abstract and secondary summary.)
15. Gyllenhaal C, Kadushin MR, Southavong B et al. (2012). Ethnobotanical approach versus random approach in the search for new bioactive compounds: support of a hypothesis. Pharmaceutical Biology 50, 30-41. https://scholars.hkbu.edu.hk/en/publications/ethnobotanical-approach-versus-random-approach-in-the-search-for–2/ (abstract only. Author list beyond the first [VERIFY].)
16. Haygarth J (1800). Of the imagination, as a cause and as a cure of disorders of the body, exemplified by fictitious tractors. Bath, R. Cruttwell. https://www.jameslindlibrary.org/haygarth-j-1800/ (James Lind Library summary and secondary accounts. Original not read.)
17. Heinrich M, Ankli A, Frei B, Weimann C, Sticher O (1998). Medicinal plants in Mexico: healers' consensus and cultural importance. Social Science and Medicine 47(11), 1859-1871. https://www.scinapse.io/papers/2136694751 (abstract and secondary summaries.)
18. Hill AB (1965). The environment and disease: association or causation? Proceedings of the Royal Society of Medicine 58(5), 295-300. https://aliquote.org/pub/Hill_1965.pdf (full text read.)
19. Hróbjartsson A, Gøtzsche PC (2001). Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. New England Journal of Medicine 344. https://chiro.org/research/ABSTRACTS/Is_the_Placebo_Powerless.shtml (abstract and secondary summaries.)
20. Hróbjartsson A, Gøtzsche PC (2010). Placebo interventions for all clinical conditions. Cochrane Database of Systematic Reviews, Issue 1, CD003974. https://pubmed.ncbi.nlm.nih.gov/20091554/ (abstract only.)
21. Hektoen International (2017). William Withering and the foxglove. https://hekint.org/2017/01/27/william-withering-and-the-foxglove/ (secondary, read via search extract.)
22. James Lind Library (undated). Withering W (1785) entry and biographical sketch. https://www.jameslindlibrary.org/?p=1177 (read in part.)
23. James Lind Library (undated). Semmelweis I (1861) entry. https://www.jameslindlibrary.org/semmelweis-i-1861/ (secondary summary via search extract.)
24. John Innes Centre (undated). Vinblastine and vincristine: life-saving drugs from a periwinkle. https://jic.ac.uk/blog/vinblastine-and-vincristine-life-saving-drugs-from-a-periwinkle/ (secondary, read via search extract with Canadian Medical Hall of Fame biographies. Date of blog [VERIFY].)
25. Kaptchuk TJ (1998). Powerful placebo: the dark side of the randomised controlled trial. Lancet 351(9117), 1722-1725 [VERIFY volume and pages]. https://sciencebasedmedicine.org/?p=15290 (original not read. Read via secondary commentary and search extracts.)
26. Kaptchuk TJ, Kelley JM, Conboy LA et al. (2008). Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. BMJ 336, 999-1003. https://www.bmj.com/content/336/7651/999 (abstract.)
27. Katch FI (undated). James Lind. Sportscience history pages. https://sportsci.org/news/history/lind/lind_sp.html (secondary summary of Lind 1753.)
28. Kienle GS, Kiene H (1997). The powerful placebo effect: fact or fiction? Journal of Clinical Epidemiology 50(12), 1311-1318. https://pubmed.ncbi.nlm.nih.gov/9449934/ (abstract and secondary summaries.)
29. Leonti M (2011). The future is written: impact of scripts on the cognition, selection, knowledge and trans-

mission of medicinal plant use and its implications for ethnobotany and ethnopharmacology. Journal of Ethnopharmacology 134, 542-555. https://iris.unica.it/handle/11584/182586 (abstract only.)

30. Leonti M, Staub PO, Cabras S, Castellanos ME, Casu L (2015). From cumulative cultural transmission to evidence-based medicine: evolution of medicinal plant knowledge in Southern Italy. Frontiers in Pharmacology 6, 207. https://www.frontiersin.org/articles/10.3389/fphar.2015.00207/pdf (abstract via search extract. Title and author list [VERIFY].)
31. McDonald CJ, Mazzuca SA, McCabe GP (1983). How much of the placebo effect is really statistical regression? Statistics in Medicine 2, 417-427. https://www.dcscience.net/Senn-McDonald-correction.pdf (secondary only, via Senn's 1988 correction letter.)
32. Moerman DE (1991). The medicinal flora of native North America: an analysis. Journal of Ethnopharmacology 31, 1-42. https://deepblue.lib.umich.edu/items/aa01b751-1ca5-4f02-b865-a680a7fb22f1 (abstract and secondary description.)
33. Moerman DE, Pemberton RW, Kiefer D, Berlin B (1999). A comparative analysis of five medicinal floras. Journal of Ethnobiology 19(1), 49-67. https://ethnobiology.org/sites/default/files/pdfs/JoE/19-1/Moerman-etal.pdf (full text read.)
34. Moerman DE (2000). Cultural variations in the placebo effect: ulcers, anxiety, and blood pressure. Medical Anthropology Quarterly 14(1), 51-72. https://moodle.brynmawr.edu/pluginfile.php/24321/mod_resource/content/1/Moerman.%20Cultural%20Variations%20in (full text read. Source for Thomas 1987 and Horwitz et al. 1990, not read directly.)
35. Moerman DE, Jonas WB (2002). Deconstructing the placebo effect and finding the meaning response. Annals of Internal Medicine 136(6), 471-476. https://chiro.org/research/ABSTRACTS/Deconstructing_the_Placebo_Effect.shtml (abstract.)
36. Moerman DE (2002). Meaning, Medicine and the "Placebo Effect". Cambridge University Press. https://assets.cambridge.org/97805218/06305/frontmatter/9780521806305_frontmatter.pdf (not read. Listed for further reading only.)
37. Morabia A (2004). Pierre-Charles-Alexandre Louis and the evaluation of bloodletting. James Lind Library Bulletin. https://jameslindlibrary.org/articles/pierre-charles-alexandre-louis-and-the-evaluation-of-bloodletting/ (full text read.)
38. Morton V, Torgerson DJ (2003). Effect of regression to the mean on decision making in health care. BMJ 326, 1083-1084. https://eprints.whiterose.ac.uk/42/1/1083.pdf (full text read.)
39. Newmaster SG, Grguric M, Shanmughanandhan D, Ramalingam S, Ragupathy S (2013). DNA barcoding detects contamination and substitution in North American herbal products. BMC Medicine 11, 222. Retracted 2024. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC11229018/ (retraction notice via search extract.)
40. Nortier JL, Martinez MC, Schmeiser HH et al. (2000). Urothelial carcinoma associated with the use of a Chinese herb (Aristolochia fangchi). New England Journal of Medicine 342, 1686-1692. https://www.reliasmedia.com/articles/51046-aristolochia-fangchi-and-urothelial-carcinoma (abstract and secondary summaries.)
41. Royal Botanic Gardens Kew (2017). State of the World's Plants 2017, press release. https://www.kew.org/about-us/press-media/state-of-the-worlds-plants-2017 (press release and secondary coverage.)
42. Saslis-Lagoudakis CH, Savolainen V, Williamson EM et al. (2012). Phylogenies reveal predictive power of traditional medicine in bioprospecting. PNAS 109(39), 15835-15840. https://pmc.ncbi.nlm.nih.gov/articles/PMC3465383 (full text read. Author list beyond the first [VERIFY].)
43. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. New England Journal of Medicine 358, 252-260. https://www.sciencedaily.com/releases/2008/01/080116202037.htm (abstract and press coverage. Volume and pages [VERIFY].)
44. UC Berkeley (2004). Samoa and UC Berkeley agreement on the prostratin gene sequence, news release.

https://newsarchive.berkeley.edu/news/media/releases/2004/09/29_samoa.shtml (read via search extract.)

45. Vickers A, Goyal N, Harland R, Rees R (1998). Do certain countries produce only positive results? A systematic review of controlled trials. Controlled Clinical Trials 19, 159-166. https://vivo.weill.cornell.edu/display/pubid9551280 (abstract.)
46. Werner D, Thuman C, Maxwell J (2024 edition). Where There Is No Doctor, chapter 1, Home cures and popular beliefs. Hesperian Health Guides. https://hesperian.org/wp-content/uploads/pdf/en_wtnd_2024/en_wtnd_2024_01.pdf (full chapter read. Co-author list for this edition [VERIFY].)
47. Evans I, Thornton H, Chalmers I, Glasziou P (2011). Testing Treatments: better research for better healthcare, second edition. Pinter and Martin. https://www.testingtreatments.org/book/about/ (website sections only. Book not read in full for this report.)
48. Trotter RT, Logan MH (1986). Informant consensus: a new approach for identifying potentially effective medicinal plants. In Etkin NL (ed.), Plants in Indigenous Medicine and Diet, 91-112. https://experts.nau.edu/en/publications/informant-consensus-a-new-approach-for-identifying-potentially-ef/ (catalogue record and secondary formula only.)