Why does the standard RCT not apply to homeopathy?
When I started to look at this problem I did not have the resources available to Prof Hahn. Instead I looked for fundamental reasons why the analysis might be wrong, to see if there was methodological bias.

The RCT itself is fairly blameless, it is simply a neat method for multiple comparisons.
Blinding and randomisation we know add to the probity of the comparisons, I have no argument with that (except when the blinding protocol is handed to a conjurer with a vested interest).

Where it may fall down is when the design is poor, where it does not take into account the reality of the process it is measuring, but is incorrectly reductionist making hidden assumptions, and to an extent when the limits of outcome are set too tightly so that things are missed.

However, the main problem is that RCTs are generally assessed by use of frequentist statistical methods.
Much is known about frequentist inference, and part of that knowledge is that the unique case is not addressed by looking at the average which is the basis if frequentism.
It is central to homeopathy that each patient is treated as a unique case.
It aims to encourage the health of the one patient rather than to kill the vector of the disease in all of them
Those who dislike homeopathy presumably do not like to hear this, so they feel free to disregard it.
At this point, they start to measure the wrong thing. To measure homeopathy properly on its own term, you cannot do this.

They cannot see why - and I hope this is not raising a straw man, I really do see this attitude in them - they cannot see why if a particular remedy will cure a disease in one patient, that it can be inappropriate to another patient with the same disease.

But that is how homeopathy works, that is what you must measure. To attempt to force it to comply with a different - wrong - model is just bad science.

Furthermore, with the remedy matched to the patient, it may cure not only the main disease which is seen as the subject of the trial, but also one or more other concomitant conditions the patient has.
Again, for those immersed in a different mechanical model of reality, this is difficult to grasp. They are expecting only one measured narrow outcome.

To add to this complication that a single patient is uniquely treated (n=1, almost always), something that homeopaths do all the time the standard RCT analysis sees these as outriders, edge cases, and medical statisticians are trained to disregard them.
This thinking may even extend to the automated program used, which may helpfully delete these data points.
A splendid cure may be entirely missed, and only few are aware that the outrider may indeed be just noise, or else your next Nobel prize.

(Of interest is that n=1 does not usually apply during epidemics.)

Recap:
Frequentist statistical inference misses the individual case.
The individual case is the basis of homeopathic method
The hidden assumption that a whole population with a "particular diagnosed condition" is uniform with respect to treatment is unwarranted. (You may think of other examples such as the effects of penicillin.)

After this the methodological bias gets even more interesting.

Such is the blind faith in this variety of RCT that many believe that a working cure *must* show up in some way as effective.
Curing a particular patient (something the patient finds quite useful) has been lost in attempting to cure the entire group, as if the individual is somehow an unwelcome phenomenon.

Scientists do remarkable things sometimes in terms of approximate redefinition, at which point common usage fall out of favour.
Thus effectiveness has been redefined from curing a number of patients, to curing a herd.
I sometimes cynically point out that this new idea of measure is of less utility to the patient than it is to a company that wants to make maximum profit out of a diagnosis in a general population.

It turns out that if one can match a different remedy (more or less) to each member of a group with a certain disease, and each of those remedies is actually effective - it promotes cure for that patient - then the standard analysis can *completely miss* the curative potential in each and every one of them.
Nowadays that would be seen as a viridical paradox, but not so long ago there were statisticians who already knew this. It's just that this competence is being lost (and we should question why that is).

I did this originally by matching names, but now I tend to demonstrate it as my "padlock RCT in padlock world", which could be seen rather grandly as a thought-experiment. There is a little whimsy, don't let it put you off.

The Padlock RCT

Consider a population of padlocks.
In padlock-world the padlocks are ideal in that they cannot be 'tricked', and have unique key each. The rightly matched key definitely works for any particular padlock.

A syndrome
Let us postulate that that in padlock-world, some padlocks suffer a syndrome of being permanently locked up, unable to enjoy their full potential.

orthodoxy
Very often an application of oil is the first recourse for treatment.
Padlocks with the syndrome look and feel better, but generally do not open. It is known not to be a cure, but is relatively inexpensive.
It has been found that an angle-grinder can be used to open a padlock in this condition, though its subsequent quality of life may suffer a little.

a reported cure
It has been reported that one of these padlocks with 'locked-up syndrome', padlock "A" has found sudden relief by using a remedy called 'a key'. This is anecdotal.
'Key' has been tried before by scientists in padlock-world. It is regarded as alternative therapy. It may work for one but not for another. Key-opathic 'doctors' choose from a bewildering variety of so-called 'keys', claiming to find a direct match to the patient. It does not seem to make sense.

The 'key' given to padlock "A" as remedy used is known to key-opaths as remedy 'a'; it is claimed to be a matching individualised holisic remedy, *which is to say it achieves cure by being matched to the patient and his(her)(it's) symptoms, and not in one-to-one relation with the disease alone.*

It is well-known by scientists in padlock-world and perfectly obvious that one disease always has one remedy; individualised remedies are mere folk medicine, unsciemtific. They crop up one case at a time, and are not worth consideration. Outriders. Edge cases. The patient probably has a vivid imagination, cannot be trusted.

However, clinicians have seen that key 'a' indisputably works for padlock "A". Very strange.

a preliminary experiment
Let's say there was an original experiment after the first anecdotal observation; four patients with locked-up syndrome {A,B,C,D} were gathered, and the one 'key' (remedy 'a'), believed to work to relieve the disease, was applied to each.

Results were highly significant for the one patient, A, but otherwise useless. The trial showed success in only 25% of the sample; also the size of the sample is very small so by convention even that result would not be not of any great statistical significance.

opportunity for profit
Scientists versed in EBM decide to test the claim.

The RCT has been a very useful tool to evaluate potentially profitable panaceas. Although it really only shows optimal profit yield, they have even managed to convince regulators that it is the best and essential way to test a novel remedy for effectiveness, to the extent of discounting actual patient experience.
It is a very convincing illusion, appearing as it does to compare like with like.

A conventional trial
Take a sample of 26 such padlocks {A, B, C....Z}, say, each with locked-up syndrome.

For the purposes of describing this thought-experiment briefly and clearly, I shall dispense with randomization, and use only a small group for placebo control. (The padlock-world experimenters themselves would have to have been blinded, lest they distort the experiment.)

In a sample of 26 'patients' including 'A', treated with 'key' (accurate copies of key 'a') only one patient (A) responds with cure.

According to this trial, it seems remedy 'key' has less than 5% chance of working, *for the specific condition,* scarcely significant.

According to this, the use of 'key' does not work.

Nevertheless, given the startling effect on one patient, it is felt that further examination is still warranted.

further trials
A further sample of 60 padlocks with 'locked-up' syndrome is co-opted. Each is tested with remedy 'key'. There are no successes.
It seems to the pharmaceutical statistician and his followers that 'key' *does not work* at all.

..Bear with me..

final trial
Nevertheless another trial of a further 100 'patients' is tried, again with no success. It seems patently clear that this so-called remedy
'a key' *does not work at all* as a cure for the specific condition; there is no profit in it, it is quackery, anecdote, folk tradition.

further analysis to confirm result
As more damning proof, we can use further frequentist statistics (fiscal statistics) to show that so-called 'key' therapy does not work.

regression to mean
We can see clear evidence of regression to mean showing that 'key' therapy is ineffective; larger trials - those higher standard, "better quality" trials, according to the convention - show worse results.

They show clearly that the keyopathic approach does not work.

Agglomerated trials, adding more patients, show worsening results are population size is is increased - a clear indication by regression to mean that keyopathy *does not work*.

meta-analysis
We can achieve a homogeneous meta-analysis which will show that without doubt, there is only tiny 'effect' for remedy 'a key' when applied to the specific condition of 'locked-up syndrome'.

funnel plot
A funnel plot also shows 'key' does not work.

4	25.0%	--	
26	3.9%	20%	5
60	0.0%	--	
100	0.0	--	

frequentist conclusion
Keyopathy is obvious quackery, complete bunk, and should not be funded. The book is closed. No more research should be permitted.

further multi-centre trials
such is the interest that further trials of similar groups are carried out at other centers of excellence
At center B, they test key b. It is shown to fail in a similar way.
At center C, they test key c. It is shown to fail in a similar way.
At center D, they test key d. It is shown to fail in a similar way.
.... and so on to center Z

"The book should be closed. No more research should be permitted."

==========
==========

It is the incorrect interpretation of scientific method that fails. One should ask which is the quack.

Incidentally, the analogy may be extended more closely to resemble homeopathy, where say 12 to 300 or more remedies might be prescribed for a certain condition. There are even further complications to the method.

One can do interesting things with the model by introducing confounding factors. since placebo is reckoned to affect five to thirty percent of cases, a remedy which *does* work for less 5% of cases can get quite lost. Incidentally, placebo does not produce permanent cure as does a correct homeopathic remedy.

The LLN still applies, of course, except that each remedy settles down to its *appropriate proportion* of patients having a particular condition, and each may well show effectiveness for only a small proportion. The large number has to be quite large., which would make regulatory testing overly onerous as well as unnecessary (considering that the basis of repertories is that remedies are reported to work)

None of this is really ground-breaking. It should be known. It will I'm sure be of increasing interest in oncology and gene-related medicine, where drugs will be shown to work for only small numbers of the while group.