MAIMON WORKING PAPERS No. 14 AUGUST 2023

THE CHALLENGE FOR HEALTH TECHNOLOGY ASSESSMENT: PAPER MILLS, FALSE CLAIMS AND THE ENDORSEMENT OF IMAGINARY CLAIMS

Paul C. Langley, Adjunct Professor, College of Pharmacy, University of Minnesota, Minneapolis MN

ABSTRACT

Health technology assessment (HTA) has the odd distinction of being a discipline, in fact the only discipline, where non-empirically evaluable claims for product value are created by assumption driven modeled simulations. This sets HTA aside from the physical sciences and other social sciences, including mainstream economics, where the consensus is that all claims should meet the standards of normal science and fundamental measurement. The value creation belief system in HTA rejects this standard; one that has been in place since the scientific revolution of the 17th century in all other disciplines. False claims in HTA arise not only from imaginary modeled simulations, which are defended because their assumptions for an unknown future are judged realistic, but because inputs to these models, the so-called realistic assumptions, are either patently false or, in the case of assumptions based on value claims from randomized clinical trials (RCTs) have a high probability of being false. The purpose of this commentary is to point out that these deficiencies can be overcome by adopting the standards of normal science: all value claims, whether for clinical or patient reported outcomes and resource allocation should be in single value terms and supported by a protocol to propose how the claim can be evaluated and reported to health system decision makers. Manufacturers should be challenged to stand behind their claims for value in health. The HTA imaginary claims meme must be rejected. Whether this occurs, with so many vested interests in imaginary claims, remains an open question.

INTRODUCTION

It is important to be clear as to the meaning of the word fraud when applied to value claims for pharmaceuticals: it is the intent and knowledge of making false or misleading statements about the benefits, efficacy, safety, or other characteristics of a product. This could involve presenting inaccurate data, manipulating results, omitting important information, distorting the truth to deceive consumers, healthcare professionals, regulatory agencies or the general public. In these terms the activities of paper mills are not only clearly fraudulent, but raise a further issue for considering activities potentially fraudulent if intended to support health system decision making: the creation by assumption of modelled value claims where the value claims are not open to empirical assessment. This is creation of evidence by design, a key activity in health technology assessment (HTA), to produce non-evaluable claims for cost-effectiveness. whether this is fraud or merely deceptive or misleading is an open question. But what is not is doubt is the opportunities actively to promote deceptive, misleading or imaginary assumption driven simulated claims [1].

The purpose of this commentary is to consider the room for abuse, not only from paper mills and fabricated assumption driven simulated claims but also the extent to which fraud is potentially prevalent in the academic community in support of value claims for pharmaceutical products. The focus is on the contribution to fraudulent behavior of assumption driven simulations to produce QALY-based non-evaluable claims for cost-effectiveness. Just as paper mills can fraudulently create for a fee what are known as zombie claims or manufactured manuscripts for a product, sophisticated enough to pass muster for peer review, so the assumption driven modeled claim can also create an equivalent zombie claim by supporting a sponsor's product with the model passing muster at peer review with guidelines promoted by the HTA leadership and leading journals.

ENHANCING FRAUDULENT INTENT

There is an obvious incentive towards fraud among professionals in health technology assessment, as in other disciplines, if it is seen as necessary for graduation and first steps on a career ladder or as an opportunity to accelerate promotion with a side bonus of peer recognition, research funds, consulting opportunities and achievement awards. The publish or perish threat is always present and the need to stay on the lifetime career treadmill becomes an everyday preoccupation. If an annual review of publications, hopefully in leading journals, is the challenge then a more concerted effort at a high productivity publication profile supported, not by your research team alone, but by willing offsiders such as paper mills, becomes an attractive proposition. Paper mills and predatory open access journals, would not exist if there was not, like drug cartels in Mexico, a demand for their services.

The willingness to seek support for an improved publication profile has been supported by the growth in what may be described as predatory journals, defined as journals which present themselves as a legitimate academic or scientific publication, but engage in unethical and deceptive practices. There are literally hundreds. Their characteristics include

- Lack of rigorous peer review
- Quick acceptance
- High publication fees
- False or dubious indexing
- Solicitation emails
- Low quality website and content
- No reputable editorial board
- No academic standards
- Plagiarism and unoriginal content
- Overemphasis on fees

One of early lists of predatory journals is Beall's List, discontinued under pressure in 2015; currently, without endorsement, is the extensive yet anonymous Predatory Reports online database [2]. The fact that the authors of this database feel it necessary to protect their identities speaks volumes about the willingness and ability, often successful, of open access predatory journals to defend their position. Most recently, as a case in point for the growing concerns for sound research practice, Clarivate, owner of the Web of Science database, earlier this year delisted 82 journals for

failing to meet its improved quality criteria. Of these journals, 15 were from the scientific publisher Hindawi (owned by John Wiley since 2021). Also included was a leading journal from the open access publisher MDPI, the *International Journal of Environmental Research and Public Health* (IJERPH) which published 9,500 papers in 2020 and 17,000 in 2022. It is worth noting that this journal appears on the Predatory Journals list. At the same time, MDPI as an open access publishing house has attracted mixed reviews. These stem from the fact that while it hosts high value journals with an overall good quality it also engages in extensive special issue recruiting with, in 2020, an average of 100 special issues for each of 74 journals rising to 500 in 2021. While this has been labelled aggressive rent seeking rather than predatory behavior, there are questions of maintaining quality through a tightly controlled review period and whether this special issue strategy is sustainable without reducing to more predatory standards to maintain growth [3].

Increasingly, government agencies and academic groups are attempting to isolate assessed predatory journals (who can easily circumvent by a name change). In Norway, for example, is the official list, The Register for Scientific Journals, Series and Publishers which consists of publication channels which are not considered to be scientific [4]. Interestingly, where publication is a criterion for funds allocation in education, the acceptance of a publication for performance evaluation, means it has to be in a listed journal.

For those who are willing to go one step further there are the paper mills. These are websites, Russia is a favored location, that sell pre-written or customized academic papers, essays, research articles or other written content available for a fee. The website may present a list of thousands of possible academic papers which can be purchased and customized by the purchaser to submit for publication as their own work. Characteristics of these papers are:

- Customized content tailored to specific subject areas, academic levels and writing styles to meet buyer requirements
- Plagiarism and unoriginal content recycled from existing sources
- Low quality writing
- Quick turnaround times for special requests
- Minimal or nonexistent research and rigor
- No ethical conduct
- Exploitative practices
- Author listing (first author the highest fee)

There is an obvious link with 'dodgy' journals. Given the likely absence of any quality assessment or ethical considerations, the product of academic paper will will have a receptive audience once 'article processing' fees are paid. Although there is no data to support this, there is the obvious incentive for paper mills to launch new journals or hijacking existing journals to ensure a willing audience for situations where the journal may be considered legitimate; which leads to a speculative proposition that if the intent is to undercut the integrity of a commitment in a country to academic excellence, then sponsored paper mills and dodgy journals provide a great opportunity.

Over recent years some effort has gone into establishing standards for journal editors to apply to individual papers, for example the activities of the Committee on Publication Ethics (COPE) [5]

COPE). In HTA, on the other hand, there are no guidelines written to exclude papers that may be considered misleading, false or fraudulent; instead, there is an active endorsement of assumption driven imaginary simulations to produce non-evaluable cost-effectiveness claims tailored to journal submission standards; the best example is the CHEERS 2022 guidance [1 6] . This journal submission guidance is open to abuse given the absence of any basis for claims assessment, together with their neglect of the standards of normal science and fundamental measurement. Indeed, claims that the guidance has been followed may be seen as a reason for not digging deeper into the model structure and role of assumptions by journal editors and peer reviewers.

A growing area of concern is the presence of paper mill products in invited special supplements to journals where academics who are linked as advisers send invitations as guest editors to colleagues to publish on a selected topic. It is not a question of challenging what may be seen as an entirely kosher activity, but to point out the substantial financial gains to publishers from special supplements where their promotion has grown exponentially over the past few years with increasing evidence of paper mill contributions and poor standards for peer review.

CLINICAL TRIALS FALSE REPORTING

The issues of predatory journals and paper mills will not go away any time soon; indeed, they look set to become a continually morphing feature of academic research reporting. The journal *Nature* has published a number of reviews over the past few years to come to grips with the magnitude of the problem [7] .

In respect of clinical trials, building on a study reported in *Anaesthesia* that evaluated 526 trial submissions from February 2017 to March 2020 together with individual level data submitted by authors of 153 trials while categorizing trials with false data as zombie trials [8], the first cut of the submissions found 14% had submitted false data with 8% categorized as zombie. Access to individual patient data increased detection of false reporting to 44% versus 2% for the balance of submissions.

In an accompanying Editorial, given the majority of submissions reviewed in *Anaesthesia* were from only 5 countries (China, South Korea, India, Japan and Egypt) the question was raised as to the possibility of extrapolating these findings from what should be considered suspect countries [9]. Given data for suspect counties where spreadsheets were routinely sought almost all of the false trial data submitted put these trials in the zombie category. Applying the country estimates to the WHO International Clinical Trials Registry (26 September 2020), extrapolation yields for 7 countries (addition of Turkey and Iran) the estimate was 90,000 registered false trials and 50,000 zombies; an understatement as registration was not compulsory. The extrapolation for these countries yielded 200,000 to 300,00 false trials and 100,000 to 200,000 zombies. Although we would expect a high proportion of quality trials from countries such as the US, UK and Australia, the overall extrapolation is swamped by the trial product in countries such as China. This does not mean we only have to raise questions of these countries, there is much poor research in the US, UK and Australia (and others). More worrying is that the level of sophistication could translate to greater sophistication in the zombie outcome: choice of clinically irrelevant outcomes, misleading

choice of comparators and non-inferiority designs, misleading yet sophisticated statistical analysis, selective reporting spin, and other forms of dissemination bias.

The presumption that as many as 50% of clinical trial submission may include false data, zombie reporting, with a proportion submitted by paper mills and allowed through by predatory journals, has a salutary lesson for HTA where the flagship product is the creation of imaginary assumption driven simulated cost-effectiveness claims. Indeed, the application of the term imaginary takes on a new meaning where the clinical data inputs to the model itself are themselves imaginary or false claims; taken at face value with no attempt to dig deeper into the veracity of the claimed clinical impact we have imaginary assumptions driving imaginary claims. Hence the importance of protocol driven replication of clinical value claims.

We still face, with the claims for analysis of real-world data or direct real world evidence from observational studies, the question of zombie claims remains. If editors are either reluctant or claim not to have the resources to evaluate randomized clinical trials, then the playing field remains wide open for predatory journals and paper mills to look to observation studies in clinical value claims as a veritable cornucopia.

A LIFETIME COMMITMENT.

Outside of targeting dodgy journals and engagement with a paper mill, there is the attraction of mainstream manipulation or creation of data over the career of a researcher. There are numerous examples with the insurance policy of retracting a paper once it receives undue attention. This, however, can take years with the author putting up roadblocks, including threats of litigation. To which we might add the reluctance of the university and even the perpetrators colleagues to rock the boat, in many instances with accusations of scaremongering. One approach to assessing this impact is through retractions published by journals. These can take time (even years to uncover) but they point to a pattern of potentially false results.

A classic case is that of Diederek Stapel, a professor of cognitive psychology in the Netherlands. Although referred to as an unprecedented case of research fraud, it is probably the tip of the iceberg as just one extreme example of activities by researchers that merit, even on a continuum of data creation and manipulation, the term fraud [10] [11]. Concerned about the source of alleged respondent level data and how those data were manipulated, a report commissioned by the three universities where he had taught, released in 2011,found evidence for fraud having been committed in some 55 of his papers published over 10 years as well as in 10 Ph.D. dissertations where fraudulent data had been provided to the student. The field of psychology was also indicted by the report where, from a New York Times investigation [12], it was found:

> *……that Stapel's fraud went undetected for so long because of "a general culture of careless, selective and uncritical handling of research and data." If Stapel was solely to blame for making stuff up, the report stated, his peers, journal editors and reviewers of the field's top journals were to blame for letting him get away with it. The committees identified several practices as "sloppy science" — misuse of statistics, ignoring of data that do not conform to a desired hypothesis and the pursuit of a compelling story no matter how scientifically unsupported it may be.*

The *New York Times* commented:

> *The adjective "sloppy" seems charitable. Several psychologists I spoke to admitted that each of these more common practices was as deliberate as any of Stapel's wholesale fabrications. Each was a choice made by the scientist every time he or she came to a fork in the road of experimental research — one way pointing to the truth, however dull and unsatisfying, and the other beckoning the researcher toward a rosier and more notable result that could be patently false or only partly true. What may be most troubling about the research culture the committees describe in their report are the plentiful opportunities and incentives for fraud. "The cookie jar was on the table without a lid" is how Stapel put it to me once. Those who suspect a colleague of fraud may be inclined to keep mum because of the potential costs of whistle-blowing.*

It is important to note that the more egregious forms of fraud relate to the invention or manipulation of respondent or patient data; the Stapel staple of fraudulent data sets that, on closer examination, proved highly suspect. The position of a number of journals where trial or study results are reported is to ask for the original spreadsheets to re-calibrate the statistical assessment for distributions and treatment effect claims; this has led to retractions as well as evidence of multiple 'readjusting' over a number in a number of papers. It is this that revealed Stapel's fraud where is became obvious that the spreadsheets were made up, including those presented to graduate students to support their theses. Exactly the same approach could be taken to evaluate the data inputs to assumption driven cost-effectiveness simulations with concerns regarding selective choice and adjustments to the original realistic data.

OPEN SEASON IN HEALTH TECHNOLOGY ASSESSMENT

As far as can be judged, there is little attention (if any) given to the potentially pernicious impact of paper mills or the embrace of dodgy journals in HTA. This is surprising when we consider the central role HTA in the creation of claims for cost-effectiveness with assumption driven simulations, in formulary decisions, pricing and access. The problem is that when approximate invented information takes center stage there is no basis for judging its merits; one modelled imaginary claim is as good as another, except for those who hold to the naïve belief that we can judge the merits of a model by the realism of its assumptions. The inputs to support assumptions may themselves be imaginary but with imaginary claims there is no incentive to revisit those assumptions.

A further point, which no-one wants to highlight, is that HTA puts the standards for fundamental measurement to one side. Rather than recognizing that value claims for competing pharmaceutical products can only be evaluated with standard statistical technique where the claim is expressed as a single attribute with unidimensional, linear, interval and invariant properties, HTA focus instead on the mathematically impossible notion of a quality adjusted life year (QALY) [13]. The QALY fails because it is impossible to discount time spent in a disease state by a preference or utility score which is not only a composite or multiattribute creation but one with only ordinal properties [14]. As far as HTA is concerned, the contribution of Rasch measurement is ignored; there is no interest in the fact that, from a measurement perspective, the Rasch model provides the necessary and sufficient means to transform subjective ordinal observations or counts into linear measures [15].

Instead, algorithms are proposed for multiattribute instruments to transform ordinal observations into a composite ordinal measure; there is no concept of measurement where application of the term implies the previous construction and maintenance of a calibrated measuring system.

The issue, therefore, is not just one of questioning the merits, if any, of assumption driven simulated cost-effectiveness claims but a more inexcusable error: the failure to recognize the requirements of a calibrated scale for subjective responses with the desired Rasch properties. It is not the obvious logical error of believing in the realism of assumptions to create believable cost-effectiveness claims that can extend in models with time frames decades ahead, but the fact that the models are worthless in the reliance on QALYs. This is seen in the complete confusion in respect of measurement theory shown by the Institute for Clinical and Economic Review (ICER):

> *As we have expressed before we (and most health economists) are confident that changes in the EQ-5D (and other multiattribute utility instruments) do have ratio properties. The EQ-5D value sets are based on time trade-off assessments (which are interval level), with preference weights assigned to different attributes. We fail to see why this should be considered an ordinal (ranked) scale. The dead state represents a natural zero point on a health related quality of life. Negative utility values on the EQ-5D scale represent states worse than dead. We do not find this lacks face validity* [16]

This is complete nonsense; but is a key element of the HTA meme. Unless we take this at face value, assumption driven modelled claims for cost-effectiveness collapse. What is not appreciated is the standard for an interval or ratio scale established by Rasch modelling: all value claims must be for a single attribute, unidimensional, linear, interval and invariant. ICER is not alone; similar confusion (and no mention of Rasch) is shown in the leading HTA textbook (now in its 4rd ed) [17]. This textbook is the best introduction to the HTA meme as a guide to the construction of imaginary cost-effectiveness claims with the endorsement of multiattribute preference and utility scores which have no value as measures. The emphasis in this textbook on the QALY fails to recognize that with composite ordinal utility and preference scores, it is a mathematically impossible construct. The implications are of interest: claims for incremental cost-unity ratios are meaningless together with blanket claims for a product's cost-effectiveness. The intriguing feature is that the required standards for creditable measurement of subjective responses were in place over 40 years before the commitment to ordinal measurement and impossible QALYs in HTA.

But practitioners in HTA are either unaware or are content to put these annoying considerations aside. This, inevitably opens the doors to paper mills and the acceptance of modelled claims, not only by predatory journals, but by leading journals in medicine and HTA. To this open season invitation, we should recognize that HTA simulation models also have a role as promotional or marketing tools where models can be adjusted to create a favorable case for a product at a price already determined by the manufacturer. After all, unless the model seems intractable in being able to be adjusted to create a desired outcome, then it will be rejected. There is evidence, even in leading journals, that the model claims are presented only when they support the sponsor's product [18].

The ability to manipulate assumption driven simulation models not only opens up the doors to paper mills and predatory journals, but to an opportunity for consultants and medical writers to jump on the bandwagon. Why run the reputational risk of being linked to a paper mill or going down market to a predatory journal, when claiming that the model follows accepted guidelines endorsed by the leaders in HTA (e.g., CHEERS 2022) is sufficient for submission and peer review.

The fact is that creating simulated imaginary claims is trivial. A standard Markov software package can support any number of simulated claims with selecting assumptions from the literature, including clinical trial value claims; ICER actually opens the floodgate for endless manipulations of ICER's own imaginary models with the release of the ICERAnalytics software platform. This, in a real sense, provides a basis for manipulation which is no different from the application of similar software by a paper mill [19]. Indeed, the opportunity is no different from selecting a simple decision model to support a clients claim. Perhaps the answer to the question of why there is no evidence for paper mill products and dodgy journals, is that the HTA simulation modeling does not need them. Even so, there is an obvious market for paper mills in HTA for simulation models, particularly for those with no interest in the effort of modeling. The cookie jar is always open.

THE ETERNAL QALY

How has the QALY maintained is position in the traditional HTA meme when, by Rasch standards (which are unique in this regard), utilities and preference scores fail the standards of normal science and measurement for subjective observations. One explanation for this, and for the amazing number of PubMed citations for the QALY and for the term cost-effectiveness is to consider HTA as effectively embracing a strong form of relativism that captures the entire HTA meme [20]. The genesis of relativism can be traced back to Wittgenstein is his belief that truth is what we choose to make of it; truth, or meaning in use, requires a social consensus not any correspondence between what we say and any evidence for an external objective reality [21]. For a relativist, who holds to the 'strong program', the content and organization of science admits only of a sociological explanation; the values and aspiration of practitioners are to be viewed in this framework [22]. The key is the notion of symmetry: all types of knowledge claims must be treated and explained in terms of sociological or psychological imperatives. It is invalid to argue that one belief system is superior to another even if there is evidence for it. In HTA, even though assumption driven simulation models are singularly deficient in terms of the criteria for normal science and fundamental measurement, the relativist would maintain that the argument is irrelevant. They have equal status with an HTA paradigm that was based on the standards of normal science and fundamental measurement. The fact that scientific arguments can be resolved, at least provisionally, by an appeal to superior evidence is of no interest even though this makes science distinctive [19].

But there is a further question: why? Despite its manifest deficiencies, why has the HTA relativist meme endured. The relativist position is that it denies that the role of science a way of coming to grips with reality; evidence is never discovered it is constructed within a social community. HTA eschews discovery of new facts in favor of constructing evidence Science for a relativist, in other words, is about rhetoric, persuasion and authority and, given the symmetry principle, this is all it can be about [19]. The HTA belief in constructing evidence and insistence on the pre-eminent role of

assumption driven simulations in the community of believers, is subscribing to the position that truth is consensus. Unfortunately, although there will be plenty of believers defending that position, they have in non-relativist terms locked themselves into a box or an analytical dead end. The HTA meme allows no escape; it rejects by rhetoric, persuasion and authority any attempt to discover new yet provisional facts and the possibility (or hope) of modifying or overturning a paradigm. Memes are self-referential and self-sustaining and, like intelligent design, admit of no basis for overturning a consensus position.

To illustrate the extent to which the meme perpetuates false belief, consider the position taken by the Institute for Clinical and Economic Review (ICER), a leading proponent of assumption driven modelled imaginary claims in the US. Challenged to justify their use of QALYs in simulation models, the ICER position (as previous noted) is that most health economists are confident that utility and preference scores are actually ratio measures; or at least are well disguised ratio measures. This is arrant nonsense and fails completely to recognize the unique contribution of the Rasch model for creating interval and ratio scales [23]. ICER's false belief is indicative of the extent to which false beliefs are held, in this case in support of their business model. There is no way ICER will ever change this position because its reputation and business case rests on these manifest deficiencies.

But we don't live in a relativist world; we live, at least in developed countries, in a world which is committed to the discovery of new yet provisional facts not the construction of imaginary evidence; the discovery of new facts is outside the HTA purview. In this respect it is no different from the paper mill in its construction of imaginary evidence. For the HTA meme the future presages an endless series of assumption driven simulations with non-evaluable claims for cost-effectiveness to be published in leading journals; a depressing prospect.

Illustrative of this embrace of the impossible QALY is the sheer number of references to the QALY in PubMed. Counts as of 21 August 2023 from 1 January 1990 are:

- QALY 24,618
- QALY and cost-effectiveness 16,590
- QALY and Markov 6,520
- QALY and simulation models 4,053
- Cost-effectiveness 98,682

These are alarming numbers when the axioms of fundamental measurement are the standard for assessing subjective responses. Typically presented as non-empirically evaluable cost-effectiveness estimates (as incremental cost-per QALYs) these are the key input to assumption driven simulations. If it difficult to comprehend how many claims have supported decisions on the allocation of therapies when no assumption driven simulated modeled claim has any pretense at meeting the standards of normal science and fundamental measurement; the notion of demarcation between science and non-science is absent. Paper mill modeled claims are in good company.

Of course, given the prevalence of QALYs and cost-effectiveness modeled claims, we have no idea of the contribution, if any, of paper mill constructs and their studies published, which have been deliberately created or adjusted to support a particular product price. It is not sufficient to say

that simulated modeled claims can be rejected out of hand, but to address the fact that they are, hopefully, not taken seriously by their developers and, presumably, formulary committees and other health system decision makers. Indeed, according to PubMed only 78 cost-effectiveness studies have ever been retracted (0.0008%) and 3 QALY studies (0.0001%) when, in fact, if rigorous standards had been employed virtually none of the cost-effectiveness and none of the QALY studies should haven accepted by the respective journals in the first place, before listing in PubMed. In common with the extrapolation that PubMed must have tens of thousands of false clinical trial claims indexed [8] ; HTA follows close behind with the acceptance of QALY claims and cost-effectiveness pricing when these are equally false.

Further evidence for the failure of the current HTA meme is in terms of multiattribute instruments which are viewed, incorrectly, as the basis for algorithms to yield utility or preference scores and hence QALYs. Focusing on the two most popular instruments the EQ-5D-3L and EQ-5D-5l we find from PubMed that, combined with keyword EuroQol they yield 11,340 responses. Again, PubMed (and the respective journals) have the problem that these two multiattribute instruments yield composite scores which have only ordinal properties; which means they cannot support claims for therapy response (other than by comparing medians and modes with non-parametric statistics). Applying these to an assumption driven simulation model to create QAYs and cost-effectiveness claims is meaningless; the term cost-effective when applied to these model outcomes has no empirically evaluable application. It is, in other words, meaningless with PubMed indexing papers, probably submitted by unsuspecting or ill aware journal editors supported by peer reviewers, who lack the skills to support the standards of normal science and Rasch measurement.

The only application for these assumption driven simulations is as a marketing tool. There is no intent that they have any empirically evaluable content. As such there is the question, one which is raised in respect of paper mills, whether they have an ethical status or should be considered sleights of hand? Are they intended to convince the more gullible that realistic assumptions necessarily produce realistic claims. After all, if NICE engages with academic reviewers and both of them give the seal of approval to the imaginary model structure and choice of assumptions, are these to be considered the most realistic simulation models? If so, are the so-called successful reference models a unique assembly of structural and parameter assumptions guaranteed to create non-evaluable cost-effectiveness claims that have a key role in pricing, formulary position and resource allocation in the health care system? NICE, it should ne noted, is in the same bind as ICER: to retreat from a position that endorses false modeled value claims for cost-effectiveness at selected prices would lead to an unpalatable, at least in the case of NICE, political position.

If this interpretation is appropriate the HTA belief system has proved to be remarkably resilient to claims that it fails the standards of normal science. The HTA meme, it certainly does not warrant to appellation of ‘paradigm’ , is a belief system that is quite clear in its rejection of hypothesis testing in favor of the production of approximate information. Perhaps is should be described as endorsing zombie information; an appellation used to describe the output claims from paper mill over-the-counter reports. If so, the HTA meme must be admired for its success as a mind-virus where success, based in large part on its transmission fidelity, makes it hard for its adherents or victims to detect. As Dawkins made clear some 50 years ago in introducing this concept of a mind

virus, or meme, the adherent is typically impelled by a deeply held conviction that something is true, right and virtuous [24]. A positive virtue that is strongly held and unshakeable; even to extent of making virtue for the lack of evidence where faith is reinforced in being certain because it is impossible

Driven by an illogical belief in the realism of assumptions to justify value claims modelled for decades into the future, HTA fails the standard for demarcation: the application of the criteria for an appeal to evidence that distinguishes science from pseudoscience. With HTA clearly in the latter camp, the question is whether the question of fraud can be applied. Clearly, whether criteria have been developed to validate imaginary claims, the application is immaterial. There is no basis in logic for choosing one set of presumed realistic assumptions over another where the notion of realism fails the simple logic of induction: the fact that past futures have resemble past pasts does not mean that future futures will resemble future pasts. Constructing an assumption driven modeled simulation for a cost-effectiveness claim with intent to defraud or at least make a favorable case for a product is, in practice, simply a reflection of choosing one set of assumptions over another and is no different from paper mill assumption driven simulation model claims.

HTA seems uniquely placed to encourage and disguise fraud in the modeled application of clinical trial results. We have known for decades that fraud is prevalent in clinical papers and the apparent willingness of authors of systematic reviews to accept what an independent assessment would label as fraud; what has now been seen as a growing epidemic of zombie papers. Paying lip service to the role of well supported and meaningful clinical trial value claims meeting standards of normal science is all too common while simultaneously engaging in fraudulent behavior. This ranges from spin in the presentation of clinical trial results to the outright fraudulent creation of paper mill results. There are all too many instances of such behavior which has escaped the attention of editors and peer reviewers in often leading journals; typically, the primary target of paper mill providers who rely on teams of professionals to create zombie claims. It is entirely beside the point that zombie claims, accepted provisionally by a leading journal where, for a fee, an aspiring researcher can have his or her named added to the author list (a place as first author is the highest fee) may lack, on closer inspection, credibility. An inspection which seldom occurs unless a red flag is raised by an independent analyst or group such Retraction Watch [25]. Even the Cochrane Collaboration, the long-touted doyen of systematic reviews, has been slow to react, although there are now guidelines in place to inform systematic reviewers of their responsibilities, there are still concerns over the conduct of systematic reviews [26 27 28].

CONCLUSIONS

The fact that databases such as PubMed are storing thousands if not tens of thousands of fake paper mill papers or their CHEERS 2022 equivalent must cast doubt on whether HTA has a future. If we believe in the scientific method and the process of discovery of new yet provisional facts, then fake papers are a major roadblock. On the data presented in an abstract and even if the full paper is accessed, we would still have only limited information on which to assess the merits of the research claims. The concern is that criteria to evaluate the likelihood of a paper mill research report are unlikely to identify more than a small proportion of the papers and even if the success rate is considered reasonable, those producing false research papers, aided and abetted by hundreds

if not more academic researchers willing to 'sell their souls' will be shooting at a moving target. The fraud will become more sophisticated; criteria will be identified and the paper mill making adjustments to meet those criteria.

This ability to evade fraud is exacerbated by the HTA meme and its relativist belief in the creation rather than the discovery of evidence. Attempts to exclude paper mill products will be virtually impossible if, in defense of an adverse decision, authors argue that the simulation model methodology has been endorsed by the leadership in HTA (e.g., CHEERS 2022). It's a no-win situation; journal editors will not take the initiative. The only solution is to reject assumption driven simulated claims entirely; it is not a question of the bad driving out the good but the fact that there was no good to start with, given the rejection some 30 years ago of the standards for normal science and measurement.

The prospects for effective remediation are slim. Too many have too much to lose, taking refuge in the well-known tactic of denial and the threat of litigation; US presidential candidates are a role model. The bottom line is that too few seem to care about the prevalence of fraud, particularly among academic colleagues in HTA, so that despite whistle blowing the effort is largely wasted. When the question of paper mills in HTA is considered, we face the fact that in its emphasis on creating non-empirically evaluable modeled claims, HTA has shot itself in the metaphorical foot. The meme or belief system that has put approximate information ahead of hypothesis testing, supported by the HTA leadership for over 30 years, make abuse straightforward as evidenced by a recent and only comprehensive review [11]. Assumption driven simulated modeled claims for cost-effectiveness and threshold pricing that comes in just below threshold are, to all intents and purposes, marketing tools; tools that are no different from paper mills. Both go to encourage formulary decisions and product uptake, supported by the option of selecting a paper mill claim to justify assumptions in the model for clinical and PRO benefits. Perhaps the key yet unanswered question for HTA practitioners and the endorsement by journal editors of CHEERS 2022 is: is anyone concerned? After all, the creation of imaginary claims can hardly be considered research that meets accepted standards; an imaginary assumption driven claim is more appropriately viewed, following Wolfgang Pauli, as one that …. *[It] is not only not right, it is not even wrong* [29].

.

REFERENCES

[1] Langley P. Nothing to Cheer About: Endorsing Imaginary Economic Evaluations and Value Claims with CHEERS 22 [version 1; peer review: 2 approved]. *F1000Research* 2022, **11**:248 (https://doi.org/10.12688/f1000research.109389.1)

[2] Predatory Reports. https://predatoryreports.org/about

[3] Crosetto P. Is MDPI a predatory publisher. April 2021. https://paolocrosetto.wordpress.com/2021/04/12/is-mdpi-a-predatory-publisher/

[4] Norway. National Board of Scholarly Publishing. National Register for Scientific Journals, Series and Publishers. https://kanalregister.hkdir.no/publiseringskanaler/OmProsedyrer.action?request_locale=en

[5] Committee on Publication Ethics (COPE) Retraction Guidelines. https://publicationethics.org/sites/default/files/retraction-guidelines-cope.pdf

[6] Husereau D, Drummond M, Augustovski F et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 22) Statement: Updated reporting guidance for health economic evaluations. *ValueHealth*. 2022;25(1):3-9 https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-021-02204-0

[7] Van Noorden. How many clinical trials can't be trusted? *Nature*. 2023;619:455-458

[8] Carlisle J. False individual patient data and zombie randomized controlled trials submitted to *Anaesthesia. Anaesthesia*. 2021;76(4):9

[9] Ioannidis J. Hundreds of thousands of zombie randomized trials circulate among us (Editorial). *Anaesthesia*. 2021. 76:444-47

[10] Wise J. Extent of Dutch psychologist's research fraud was 'unprecedented'. *BMJ*. 2011;343:d7201

[11] Vogel G. Psychologist accused of fraud on 'astonishing scale'. *Science.* 2011;314:579

[12] Bhattacharjee Y. The Mind of a Con Man. *New York Times Magazine*. April 26, 2013 https://www.nytimes.com/2013/04/28/magazine/diederik-stapels-audacious-academic-fraud.html

[13] Langley PC and McKenna SP. Measurement, modeling and QALYs [version 1; peer review: 2 approved] *F1000Research* 2020, 9:1048 https://doi.org/10.12688/f1000research.25039.1

[14] McKenna S, Heaney A. Composite Outcome measurement in clinical research: The triumph of illusion over reality? *J Med Econ*. 2020;23(10): 1196-1204 https://www.tandfonline.com/doi/pdf/10.1080/13696998.2020.1797755?needAccess=true

[15]Wright B, Linacre J. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil.* 1989; 70(12):857-60 https://www.researchgate.net/publication/20338407_Observations_are_always_ordinal_measurements_however_must_be_interval/link/5563b02408ae9963a11ef326/download

[16] Institute for Clinical and Economic Review. Lupus Nephritis. ICER-lupus nephritis-comments-folio_031221.pdf

[17] Drummond M, Sculpher M, Claxton K et al. Methods for the Economic Evaluation of Health Care Programmes. New York: Oxford University Press, 2015

[18] Xie F, Zhou T. Industry sponsored bias in cost-effectiveness analysis: registry-based analysis. *BMJ.* 2022:377 https://www.bmj.com/content/bmj/377/bmj-2021-069573.full.pdf

[19] Langley P. Let a Thousand Models Bloom: ICER Analytics Opens the Floodgates to Cloud Pseudoscience. *Inov Pharm*. 2021;12(1):No. 5 https://pubs.lib.umn.edu/index.php/innovations/article/view/3606

[20] Wootton D. The Invention of Science: A new history of the scientific revolution. New York: HarperCollins, 2015

[21] Macgee B. Popper, London: Fontana, 1974

[22] Bloor D. Wittgenstein: A Social Theory of Knowledge. London: Macmillan, 1983

[23] Langley P. Peter Rabbit is a Badger in Disguise: Deconstructing the Belief System of the Institute for Clinical and Economic Review in Health Technology Assessment. *InovPharm*. 2021; 12(2): No.20 https://pubs.lib.umn.edu/index.php/innovations/article/view/3992/2855

[24] Dawkins R. A Devil's Chaplain. New York: HoughtonMiflin, Mariner Edition, 2004

[25] Retraction Watch. https://retractionwatch.com/

[26] Boughton S, Wilkinson J, Bero L. When beauty is but skin deep: Dealing with problematic studies in systematic reviews (Editorial). *Cochrane Database of Systematic Reviews*. 2021;(6): ED000152 https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.ED000152/full

[27] Cochrane Database of Systematic Reviews. Policy for managing potentially problematic studies: implementation guidance. May 2023. https://documentation.cochrane.org/display/EPPR/Policy+for+managing+potentially+problematic+studies%3A+implementation+guidance

[28] Saleh N. Cochrane Reviews controversy: Are the concerns valid? *Public Health and Health Policy*. February 16, 2022 https://www.mdlinx.com/article/cochrane-reviews-controversy-are-the-concerns-valid/7tjNVFB6sLR2l9VoaXqQDc

[29]https://www.azquotes.com/quote/1351389