JSLS

Limitations in Medical Research: Recognition, Influence, and Warning

Douglas E. Ott, MD, MBA

ABSTRACT

Background: As the number of limitations increases in a medical research article, their consequences multiply and the validity of findings decreases. How often do limitations occur in a medical article? What are the implications of limitation interaction? How often are the conclusions hedged in their explanation?

Objective: To identify the number, type, and frequency of limitations and words used to describe conclusion(s) in medical research articles.

Methods: Search, analysis, and evaluation of open access research articles from 2021 and 2022 from the *Journal of the Society of Laparoscopic and Robotic Surgery* and 2022 *Surgical Endoscopy* for type(s) of limitation(s) admitted to by author(s) and the number of times they occurred. Limitations not admitted to were found, obvious, and not claimed. An automated text analysis was performed for hedging words in conclusion statements. A limitation index score is proposed to gauge the validity of statements and conclusions as the number of limitations increases.

Results: A total of 298 articles were reviewed and analyzed, finding 1,764 limitations. Four articles had no limitations. The average was between 3.7% and 6.9% per article. Hedging, weasel words and words of estimative probability description was found in 95.6% of the conclusions.

Conclusions: Limitations and their number matter. The greater the number of limitations and ramifications of their effects, the more outcomes and conclusions are affected. Wording ambiguity using hedging or weasel words shows that limitations affect the uncertainty of claims. The limitation index scoring method shows the diminished validity of finding(s) and conclusion(s).

Key Words: Bias, Hedging, Limitations, Methods, Research, Uncertainty, Validity.

INTRODUCTION

As the number of limitations in a medical research article increases, does their influence have a more significant effect than each one considered separately, making the findings and conclusions less reliable and valid? Limitations are known variables that influence data collection and findings and compromise outcomes, conclusions, and inferences. A large body of work recognizes the effect(s) and consequence(s) of limitations.[1-77] Other than the ones known to the author(s), unknown and unrecognized limitations influence research credibility. This study and analysis aim to determine how frequently and what limitations are found in peer-reviewed open-access medical articles for laparoscopic/endoscopic surgeons.

This research is about limitations, how often they occur and explained and/or justified. Failure to disclose limitations in medical writing limits proper decision-making and understanding of the material presented. All articles have limitations and constraints. Not acknowledging limitations is a lack of candor, ignorance, or a deliberate omission. To reduce the suspicion of invalid conclusions limitations and their effects must be acknowledged and explained. This allows for a clearer more focused assessment of the article's subject matter without explaining its findings and conclusions using hedging and words of estimative probability.[78,79]

METHODS

An evaluation of open access research/meta-analysis/case series/methodologies/review articles published in the *Journal of the Society of Laparoendoscopic and Robotic*

Table 1.
The Ten Most Frequent Limitations Found in *JSLS* and *Surg Endosc* Articles

JSLS top 10 limitations	Total number of limitations	Number of articles	Percent of total number of limitations	*Surg Endosc* top 10 limitations	Total number of limitations	Number of articles	Percent of total number of limitations
Results not generalizable	33	33/508	6.5%	Results not generalizable	86	86/1256	6.8%
Retrospective study	32	32/508	6.3%	Selection bias	83	83/1256	6.6%
Small sample size	32	32/508	6.3%	Confounding variables and comorbidities	72	72/1256	5.7%
Confounding variables and comorbidities	23	23/508	4.5%	Retrospective study	69	69/1256	5.5%
Selection bias	21	21/508	4.1%	Small sample size	63	63/1256	5.0%
Incomplete data	20	20/508	3.9%	Incomplete data	58	58/1256	4.6%
Limited patient selection criteria	16	16/508	3.1%	Lack of standardized treatment	55	55/1256	4.4%
Limited data availability	16	16/508	5.1%	Measurement problems	53	53/1256	4.2%
No long-term follow-up	15	15/508	3.0%	Limited analysis	47	47/1256	3.7%
Reporting errors	14	14/508	2.8%	Problems with study design	39	39/1259	3.1%
	222	222/508	**43.7%**		**625**	625/1256	**49.8%**

Surgery (*JSLS*) for 2021 and 2022 (129) and commentary/guidelines/new technology/practice guidelines/review/SAGES Masters Program articles in *Surgical Endoscopy* (*Surg Endosc*) for 2022 (169) totaling 298 were read and evaluated by automated text analysis for limitations admitted to by the paper's authors using such words as "limitations," "limits," "shortcomings," "inadequacies," "flaws," "weaknesses," "constraints," "deficiencies," "problems," and "drawbacks" in the search. Limitations not mentioned were found by reading the paper and assigning type and frequency. The number of hedging and weasel words used to describe the conclusion or validate findings was determined by reading the article and adding them up.

RESULTS

For *JSLS*, there were 129 articles having 63 different types of limitations. Authors claimed 476, and an additional 32 were found within the article, totaling 508 limitations (93.7% admitted to and 6.3% discovered that were not mentioned). This was a 3.9 limitation average per article. No article said it was free of limitations. The ten most

frequent limitations and their rate of occurrence are in **Table 1**. The total number of limitations, frequency, and visual depictions are seen in **Figures 1A** and **1B**.

There were 169 articles for *Surg Endosc*, with 78 different named limitations the authors claimed for a total of 1,162. An additional 94 limitations were found in the articles, totaling 1,256, or 7.4 per article. The authors explicitly stated 92.5% of the limitations, and an additional 7.5% of additional limitations were found within the article. Five claimed zero limitations (5/169 = 3%). The ten most frequent limitations and their rate of occurrence are in **Table 1**. The total number of limitations and frequency is shown in **Figures 1A** and **1B**.
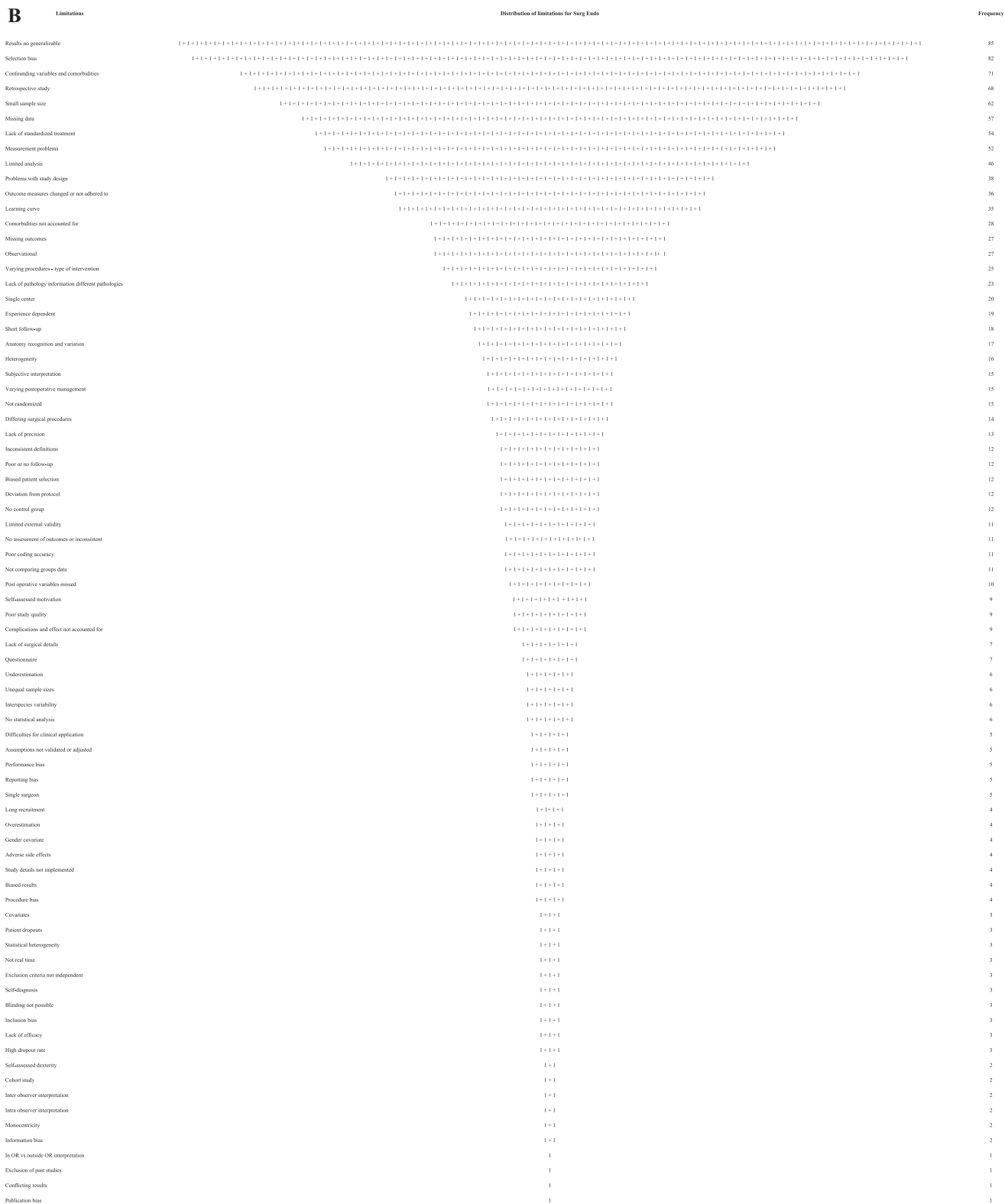
Conclusions were described in hedged, weasel words or words of estimative probability 95.6% of the time (285/298).

DISCUSSION

A research hypothesis aims to test the idea about expected relationships between variables or to explain an

A

Limitations	Distribution of limitations for JSLS	Frequency
Results not generalizable	1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1	33
Retrospective study	1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1	32
Small sample size	1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1	32
Confounding variables and comorbidities	1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1	23
Selection bias	1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1	21
Incomplete data	1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1	20
Limited patient selection criteria	1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1	16
Limited data availability	1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1	16
No long-term follow-up	1+1+1+1+1+1+1+1+1+1+1+1+1+1+1	15
Reporting errors	1+1+1+1+1+1+1+1+1+1+1+1+1+1	14
Over or under reporting	1+1+1+1+1+1+1+1+1+1+1+1+1+1	14
Technique differences	1+1+1+1+1+1+1+1+1+1+1+1+1	13
Lack of standardized assessment	1+1+1+1+1+1+1+1+1+1+1+1	12
Limited randomization comparisons	1+1+1+1+1+1+1+1+1+1+1	11
Methodologic design flaw	1+1+1+1+1+1+1+1+1+1+1	11
Single institution	1+1+1+1+1+1+1+1+1+1	10
No validated patient selection criteria	1+1+1+1+1+1+1+1+1+1	10
Qualitative assessment lacking	1+1+1+1+1+1+1+1+1	9
Recall bias	1+1+1+1+1+1+1+1+1	9
Limited full evaluation (scope)	1+1+1+1+1+1+1+1	8
No assessment criteria	1+1+1+1+1+1+1+1	8
Self-reported by patient	1+1+1+1+1+1+1+1	8
Database inconsistencies	1+1+1+1+1+1+1+1	8
Data cannot be used to calculate the specific or complete incidence	1+1+1+1+1+1+1	7
Subjective data	1+1+1+1+1+1+1	7
Different pathologies assessed	1+1+1+1+1+1+1	7
Single surgeon	1+1+1+1+1+1	6
Different ways of coding	1+1+1+1+1+1	6
Study finding incidence not representative	1+1+1+1+1	5
Follow-up was patients own assessment	1+1+1+1+1	5
Patients lost to follow-up	1+1+1+1+1	5
No comparative controls	1+1+1+1+1	5
Rate of recurrence not validated	1+1+1+1	4
Patients lost to follow-up	1+1+1+1	4
Learning curve	1+1+1+1	4
Not reaching statistical significance	1+1+1+1	4
No quality control	1+1+1+1	4
Variable experiences	1+1+1+1	4
Observational study	1+1+1+1	4
Heterogeneous study group	1+1+1	3
Few samples tested	1+1+1	3
Cost analysis may not be reproducible	1+1+1	3
Surgeon selection bias	1+1+1	3
No conclusions can be drawn	1+1+1	3
Context and culture	1+1+1	3
Limited scope of comparison	1+1+1	3
Objective data not recorded for analysis	1+1+1	3
Lack of homogeneity of meta-analysis and reviews	1+1	2
Restricted options	1+1	2
No causal correlations	1+1	2
Treatment compliance not validated	1+1	2
Surgical conversions not adjusted	1+1	2
Results not reproducible	1+1	2
Study not blinded	1+1	2
Cross-sectional study	1+1	2
One time experience	1	1
Only postoperative questions asked	1	1
Type 11 error	1	1
Case matching statistical analysis	1	1
Bias correction by regression analysis	1	1
Misspecification bias	1	1
No independent assessment	1	1
	1	1

Figure 1. (**A**) Visual depiction of the ranked frequency of limitations for *JSLS* articles reviewed.

B

Limitations	Distribution of limitations for Surg Endo	Frequency
Results no generalizable		85
Selection bias		82
Confounding variables and comorbidities		71
Retrospective study		68
Small sample size		62
Missing data		57
Lack of standardized treatment		54
Measurement problems		52
Limited analysis		46
Problems with study design		38
Outcome measures changed or not adhered to		36
Learning curve		35
Comorbidities not accounted for		28
Missing outcomes		27
Observational		27
Varying procedures - type of intervention		25
Lack of pathology information different pathologies		23
Single center		20
Experience dependent		19
Short follow-up		18
Anatomy recognition and variation		17
Heterogeneity		16
Subjective interpretation		15
Varying postoperative management		15
Not randomized		15
Differing surgical procedures		14
Lack of precision		13
Inconsistent definitions		12
Poor or no follow-up		12
Biased patient selection		12
Deviation from protocol		12
No control group		12
Limited external validity		11
No assessment of outcomes or inconsistent		11
Poor coding accuracy		11
Not comparing groups data		11
Post operative variables missed		10
Self-assessed motivation		9
Poor study quality		9
Complications and effect not accounted for		9
Lack of surgical details		7
Questionnaire		7
Underestimation		6
Unequal sample sizes		6
Interspecies variability		6
No statistical analysis		6
Difficulties for clinical application		5
Assumptions not validated or adjusted		5
Performance bias		5
Reporting bias		5
Single surgeon		5
Long recruitment		4
Overestimation		4
Gender covariate		4
Adverse side effects		4
Study details not implemented		4
Biased results		4
Procedure bias		4
Covariates		3
Patient dropouts		3
Statistical heterogeneity		3
Not real time		3
Exclusion criteria not independent		3
Self-diagnosis		3
Blinding not possible		3
Inclusion bias		3
Lack of efficacy		3
High dropout rate		3
Self-assessed dexterity		2
Cohort study		2
Inter observer interpretation		2
Intra observer interpretation		2
Monocentricity		2
Information bias		2
In OR vs outside OR interpretation		1
Exclusion of past studies		1
Conflicting results		1
Publication bias		1

Figure 1. *Continued.* (**B**) Visual depiction of the ranked frequency of limitations for *Surg Endosc* articles reviewed.

Table 2.
Limitations of Known and Unknowns as They Apply to Limitations

Knowns	**Known Knowns** Things we are aware of and understand.	**Known Unknowns** Things we are aware of but don't understand.
Unknowns	**Unknown Knowns** Things we understand but are not aware of.	**Unknown Unknowns** Things we are neither aware of nor understand.
	Knowns	**Unknowns**

occurrence. The assessment of a hypothesis with limitations embedded in the method reaches a conclusion that is inherently flawed. What is compromised by the limitation(s)? The result is an inferential study in the presence of uncertainty. As the number of limitations increases, the validity of information decreases due to the proliferation of uncertain information. Information gathered and conclusions made in the presence of limitations can be functionally unsound. Hypothesis testing of spurious conditions with limitations and then claiming a conclusion is not a reliable method for generating factual evidence. The authors' reliance on limitation gathered "evidence" data and asserting that this is valid is spurious reasoning. The bridge between theory and evidence is not through limitations that unquestionably accept findings. A range of conclusion possibilities exists being some percent closer to either more correct or incorrect. Relying on leveraging the pursuit of "fact" in the presence of limitations as the safeguard is akin to the fox watching the hen house. Acknowledgment of the uncertainty limitations create in research and discounting the finding's reliability would give more credibility to the effort. Shortcomings and widespread misuses of research limitation justifications make findings suspect and falsely justified in many instances.

The *JSLS* instructions to authors say that in the discussion section of the paper the author(s) must "Comment on any methodological weaknesses of the study" (http://jsls.sls.org/guidelines-for-authors/). In their instructions for authors, *Surg Endosc* says that in the discussion of the paper, "A paragraph discussing study limitations is required" (https://www.springer.com/journal/464/submission-guidelines). A comment for a written article about a limitation should express an opinion or reaction. A paragraph discussing limitations, especially, if there is more than one, requires just that: a paragraph and discussion. These requirements were not met or enforced by *JSLS* 86% (111/129) of the time and 92.3% (156/169) for *Surg Endosc*. This is an error in peer reviewing, not adhering to established research publication best practices, and the journals needing to adhere to

their guidelines. The International Committee of Medical Journal Editors, uniform requirements for manuscripts recommends that authors "State the limitations of your study, and explore the implications of your findings for future research and for clinical practice or policy. Discuss the influence or association of variables, such as sex and/or gender, on your findings, where appropriate, and the limitations of the data." It also says, "describe new or substantially modified methods, give reasons for using them, and evaluate their limitations" and "Include in the Discussion section the implications of the findings and their limitations, including implications for future research" and "give references to established methods, including statistical methods (see below); provide references and brief descriptions for methods that have been published but are not well known; describe new or substantially modified methods, give reasons for using them, and evaluate their limitations."[65] "Reporting guidelines (e.g., CONSORT,[1] ARRIVE[2]) have been proposed to promote the transparency and accuracy of reporting for biomedical studies, and they often include discussion of limitations as a checklist item. Although such guidelines have been endorsed by high-profile biomedical journals, and compliance with them is associated with improved reporting quality,[3] adherence remains suboptimal."[4,5]

Limitations start in the methodologic design phase of research. They require troubleshooting evaluations from the start to consider what limitations exist, what is known and unknown, where, and how to overcome them, and how they will affect the reasonableness and assessment of possible conclusions. A named limitation represents a category with numerous components. Each factor has a unique effect on findings and collectively influences conclusion assessment. Even a single limitation can compromise the study's implementation and adversely influence research parameters, resulting in diminished value of the findings, outcomes, and conclusions. This becomes more problematic as the number of limitations and their components increase. Any limitation influences a research paper. It is unknown how much and to what extent any

limitation affects other limitations, but it does create a cascading domino effect of ever-increasing interactions that compromise findings and conclusions. Considering "research" as a system, it has sensitivity and initial conditions (methodology, data collection, analysis, etc.). The slightest alteration of a study due to limitations can profoundly impact all aspects of the study. The presence and influence of limitations introduce a range of unpredictable influences on findings, results, and conclusions.

Researchers and readers need to pay attention to and discount the effects limitations have on the validity of findings. Richard Feynman said in "Cargo cult science" "the first principle is that you must not fool yourself and you are the easiest person to fool."[73] We strongly believe our own nonsense or wrong-headed reasoning. Buddhist philosophers say we are attached to our ignorance. Researchers are not critical enough about how they fool themselves regarding their findings with known limitations and then pass them on to readers. The competence of findings with known limitations results in suspect conclusions.

Authors should not ask for dismissal, disregard, or indulgence of their limitations. They should be thoughtful and reflective about the implications and uncertainty the limitations create[67]; their uncertainties, blind spots, and impact on the research's relevance. A meaningful presentation of study limitations should describe the limitation, explain its effect, provide possible alternative approaches, and describe steps taken to mitigate the limitation. This was largely absent from the articles reviewed.

Authors use synonyms and phrases describing limitations that hide, deflect, downplay, and divert attention from them, i.e., some drawbacks of the study are …, weaknesses of the study are…, shortcomings are…, and disadvantages of the study are…. They then say their finding(s) lack(s) generalizability, meaning the findings only apply to the study participants or that care, sometimes extreme, must be taken in interpreting the results. Which limitation components are they referring to? Are the authors aware of the extent of their limitations, or are they using convenient phrases to highlight the existence of limitations without detailing their defects?

Limitations negatively weigh on both data and conclusions yet no literature exists to provide a quantifiable measure of this effect. The only acknowledgment is that limitations affect research data and conclusions. The adverse effects of limitations are both specific and contextual to each research article and is part of the parameters that affect research. All the limitations are expressed in words, excuses, and a litany of mea culpas asking for forgiveness and without explaining the extent or magnitude of their impact. It is left to the writer and reader to figure out. It is not known what value writers put on their limitations in the 298 articles reviewed from *JSLS* and *Surg Endosc*. Listing limitations without comment and effect on the findings and conclusions is a compromising red flag. Therefore, a limitation scoring method was developed and is proposed to assess the level of suspicion generated by the number of limitations.

It is doubtful that a medical research article is so well designed and executed that there are no limitations. This is doubtful since there are unknown unknowns. This study showed that authors need to acknowledge all the limitations when they are known. They acknowledge the ones they know but do not consider other possibilities. There are the known known limitations; the ones the author(s) are aware of and can be measured, some explained, most not. The known unknowns: limitations authors are aware of but cannot explain or quantify. The unknown unknown limitations: the ones authors are not aware of and have unknown influence(s), i.e., the things they do not know they do not know. These are blind spots (not knowing what they do not know or black swan events). And the unknown knowns; the limitations authors may be aware of but have not disclosed, thoroughly reported, understood, or addressed. They are unexpected and not considered. See **Table 2**.[74]

It is possible that authors did not identify, want to identify, or acknowledge potential limitations or were unaware of what limitations existed. Cumulative complexity is the result of the presence of multiple limitations because of the accumulation and interaction of limitations and their components. Just mentioning a limitation category and not the specific parts that are the limitation(s) is not enough. Authors telling readers of their known research limitations is a caution to discount the findings and conclusions. At what point does the caution for each limitation, its ramifications, and consequences become a warning? When does the piling up of mistakes, bad and missing data, biases, small sample size, lack of generalizability, confounding factors, etc., reach a point when the findings becomes uninterpretable and meaningless? "Caution" indicates a level of potential hazard; a warning is more dire and consequential. Authors use the word "caution" not "warning" to describe their conclusions. There is a point when the number of limitations and their cumulative effects surpasses the point where a caution statement is no longer applicable, and a warning statement is required. This is the reason for establishing a limitations risk score.

Limitations put medical research articles at risk. The accumulation of limitations (variables having additional limitation components) are gaps and flaws diluting the probability of validity. There is currently no assessment method for evaluating the effect(s) of limitations on research outcomes other than awareness that there is an effect. Authors make statements warning that their results may not be reliable or generalizable, and need more research and larger numbers. Just because the weight effect of any given limitation is not known, explained, or how it discounts findings does not negate a causation effect on data, its analysis, and conclusions. Limitation variables and the ramifications of their effects have consequences. The relationship is not zero effect and accumulates with each added limitation.

As a result of this research, a limitation index score (LIS) system and assessment tool were developed. This limitation risk assessment tool gives a scores assessment of the relative validity of conclusions in a medical article having limitations. The adoption of the LIS scoring assessment tool for authors, researchers, editors, reviewers, and readers is a step toward understanding the effects of limitations and their causal relationships to findings and conclusions. The objective is cleaner, tighter methodologies, and better data assessment, to achieve more reliable findings. Adjustments to research conclusions in the presence of limitations are necessary. The degree of modification depends on context. The cumulative effect of this burden must be acknowledged by a tangible reduction and questioning of the legitimacy of statements made under these circumstances. The description calculating the LIS score is detailed in **Appendix 1**.

A limitation word or phrase is not one limitation; it is a group of limitations under the heading of that word or phrase having many additional possible components just as an individual named influence. For instance, when an admission of selection bias is noted, the authors do not explain if it was an exclusion criterion, self-selection, non-responsiveness, lost to follow-up, recruitment error, how it affects external validity, lack of randomization, etc., or any of the least 263 types of known biases causing systematic distortions of the truth whether unintentional or wanton.[40,76] Which forms of selection bias are they identifying?[63] Limitations have branches that introduce additional limitations influencing the study's ability to reach a useful conclusion. Authors rarely tell you the effect consequences and extent limitations have on their study, findings, and conclusions.

This is a sample of limitations and a few of their component variables under the rubric of a single word or phrase. See **Table 3**.

Limitations rarely occur alone. If you see one there are many you do not see or appreciate. Limitations components interact with their own and other limitations, leading to complex connections interacting and discounting the reliability of findings. By how much is context dependent: but it is not zero. Limitations are variables influencing outcomes. As the number of limitations increases, the reliability of the conclusions decreases. How many variables (limitations) does it take to nullify the claims of the findings? The weight and influence of each limitation, its aggregate components, and interconnectedness have an unknown magnitude and effect. The result is a

Table 3.
A Limitation Word or Phrase is a Limitation Having Additional Components That Are Additional Limitations. When an Author Uses the Limitation Composite Word or Phrase, They Leave out Which One of Its Components is Contributory to the Research Limitations. Each Limitation Interacts with Other Limitations, Creating a Cluster of Cross Complexities of Data, Findings, and Conclusions That Are Tainted and Negatively Affect Findings and Conclusions

Small Sample Size	Retrospective Study	Selection Bias
Low statistical power	Missing information	Affects internal validity
Estimates not reliable	Recall bias	Nonrandom selection
Prone to biased samples	Observer bias	Leads to confounding
Not generalizable	Misclassification bias	Not generalizable
Prone to false negative error	Observer bias	Inaccurate relation to variables
Prone to false positive error	Evidence less robust than prospective study	Observer bias
Sampling error	Missing data	Sampling bias
	Confounding factors	Volunteer bias
	Selection bias	Survivorship bias

disorderly concoction of hearsay explanations. **Table 4** is an example of just two single explanation limitations and some of their components illustrating the complex compounding of their effects on each other.

The novelty of this paper on limitations in medical science is not the identification of research article limitations or their number or frequency; it is the recognition of the multiplier effect(s) limitations and the influence they have on diminishing any conclusion(s) the paper makes. It is possible that limitations contribute to the inability of studies to replicate and why so many are one-time occurrences. Therefore, the generalizability statement that should be given to all readers is BEWARE THERE IS A REDUCTION EFFECT ON THE CONCLUSIONS IN THIS ARTICLE BECAUSE OF ITS LIMITATIONS.

Journals accept studies done with too many limitations, creating forking path situations resulting in an enormous number of possible associations of individual data points as multiple comparisons.[79] The result is confusion, a muddled mess caused by interactions of limitations undermining the ability to make valid inferences. Authors know and acknowledge but rarely explain them or their influence. They also use incomplete and biased databases, biased methods, small sample sizes, and not eliminating confounders, etc., but persist in doing research with these circumstances. Why is that? Is it because when limitations are acknowledged, authors feel justified in their conclusions? It wasn't my poor research design; it was the limitation(s). How do peer reviewers score and analyze these papers without a method to discount the findings and conclusions in the presence of limitations? What are the calculus editors use to justify papers with multiple limitations, reaching compromised or spurious conclusions? How much caution or warning should a journal say must be taken in interpreting article results? How much? Which results? When? Under what circumstance(s)?

Table 4.
An Example of Interactions between Only Two Limitations and Some of Their Components Causes 16 Interactions

Retrospective Study	Small Sample Size
Confounders	Low statistical power
Not randomized	Findings not generalizable
Inaccurate records	Large margin of error
Missing data	Biased sample

Since a critical component of research is its limitations, the quality and rigor of research are largely defined by,[75] these constraints making it imperative that limitations be exposed and explained. All studies have limitations admitted to or not, and these limitations influence outcomes and conclusions. Unfortunately, they are given insufficient attention, accompanied by feeble excuses, but they all matter. The degrees of freedom of each limitation influence every other limitation, magnifying their ramifications and confusion. Limitations of a scientific article must put the findings in context so the reader can judge the validity and strength of the conclusions. While authors acknowledge the limitations of their study, they influence its legitimacy.

Not only are limitations not properly acknowledged in the scientific literature,[8] but their implications, magnitude, and how they affect a conclusion are not explained or appreciated. Authors work at claiming their work and methods "overcome," "avoid," or "circumvent" limitations. Limitations are explained away as "Failure to prove a difference does not prove lack of a difference."[60] Sample size, bias, confounders, bad data, etc. are not what they seem and do not sully the results. The implication is "trust me." But that's not science. Limitations create cognitive distortions and framing (misperception of reality) for the authors and readers. Data in studies with limitations is data having limitations. It was real but tainted.

Limitations are not a trivial aspect of research. It is a tangible something, positive or negative, put into a data set to be analyzed and used to reach a conclusion. How did these extra somethings, known unknowns, not knowns, and unknown knowns, affect the validity of the data set and conclusions? Research presented with the vagaries of explicit limitations is intensified by additional limitations and their component effects on top of the first limitations, quickly diluting any conclusion making its dependability questionable.

This study's analysis of limitations in medical articles averaged 3.9% per article for *JSLS* and 7.4% for *Surg Endosc.* Authors admit to some and are aware of limitations, but not all of them and discount or leave out others. Limitations were often presented with misleading and hedging language. Authors do not give weight or suggest the percent discount limitations have on the reliance of conclusion(s). Since limitations influence findings, reliability, generalizability, and validity without knowing the magnitude of each and their context, the best that can be said about the conclusions is that they are specific to the study described, context-driven, and suspect.

Limitations mean something is missing, added, incorrect, unseen, unaware of, fabricated, or unknown; circumstances that confuse, confound, and compromise findings and information to the extent that a notice is necessary. All medical articles should have this statement, "Any conclusion drawn from this medical study should be interpreted considering its limitations. Readers should exercise caution, use critical judgement, and consult other sources before accepting these findings. Findings may not be generalizable regardless of sample size, composition, representative data points, and subject groups. Methodologic, analytic, and data collection may have introduced biases or limitations that can affect the accuracy of the results. Controlling for confounding variables, known and unknown, may have influenced the data and/or observations. The accuracy and completeness of the data used to draw a conclusion may not be reliable. The study was specific to time, place, persons, and prevailing circumstances. The weight of each of these factors is unknown to us. Their effect may be limited or compounded and diminish the validity of the proposed conclusions."

This study and findings are limited and constrained by the limitations of the articles reviewed. They have known and unknown limitations not accounted for, missing data, small sample size, incongruous populations, internal and external validity concerns, confounders, and more. See **Tables 2** and **3**. Some of these are correctible by the author's awareness of the consequences of limitations, making plans to address them in the methodology phase of hypothesis assessment and performance of the research to diminish their effects.

CONCLUSION

Limitations in research articles are expected, but they can be reduced in their effect so that conclusions are closer to being valid. Limitations introduce elements of ignorance and suspicion. They need to be explained so their influence on the believability of the study and its conclusions is closer to meeting construct, content, face, and criterion validity. As the number of limitations increases, common sense, skepticism, study component acceptability, and understanding the ramifications of each limitation are necessary to accept, discount, or reject the author's findings. As the number of hedging and weasel words used to explain conclusion(s) increases, believability decreases, and raises suspicion regarding claims. Establishing a systematic limitation scoring index limitations for authors, editors, reviewers, and readers and recognizing their

cumulative effects will result in a clearer understanding of research content and legitimacy.

References:

All references have been archived at https://archive.org/web/

1. Schulz K, Altman D, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332.

2. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol*. 2010;8(6):e1000412.

3. Kane R, Kane RL, Wang J, Garrard J. Reporting in randomized clinical trials improved after adoption of the CONSORT statement. *J Clin Epidemiol*. 2007;60(3):241–249.

4. Turner L, Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Syst Rev*. 2012;1:60.

5. Kilicoglu H, Rosemblat G, Malicki M, ter Riet G. Automatic recognition of self-acknowledged limitations in clinical research literature. *J AM Med Inform Assoc*. 2018;25(7):855–861.

6. Ross P, Zaidi N. Limited by our limitations. *Perspect Med Educ*. 2019;8(4):261–264.

7. Ioannidis J. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124.

8. Ioannidis J. Limitations are not properly acknowledged in the scientific literature. *J Clin Epidemiol*. 2007;60(4):324–329.

9. Puhan MA, Akl EA, Bryant D, Xie F, Apolone G, ter Riet G. Discussing study limitations in reports of biomedical studies-the need for more transparency. *Health Qual Life Outcomes*. 2012;10:23.

10. Greener S. Research limitations: the need for honesty and common sense. *Interactive Learning in Environments*. 2018; 26(5):567–568.

11. Puhan MA, Heller N, Joleska I, et al. Acknowledging limitations in biomedical studies: the ALIBI study. *The Sixth International Congress on Peer Review and Biomedical Publication* Vancouver, Canada: JAMA and BMJ, 2009

12. Guyatt G, Oxman A, Kunz R, et al.; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Bmj*. 2008;336(7650):924–926.

13. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ; GRADE Working Group. What is "quality of evidence" and why is it important to clinicians? *Bmj*. 2008; 336(7651):995–998.

14. Imrey P. Limitations of meta-analyses of studies with high heterogeneity. *JAMA Netw Open.* 2020;3(1):e1919325.

15. Schroll J, Moustgaard R, Gøtzsche P. Dealing with substantial heterogeneity in Cochrane reviews: cross-sectional study. *BMC Med Res Methodol.* 2011;11:22.

16. Alba A, Alexander P, Chang J, MacIsaac J, DeFry S, Guyatt G. High statistical heterogeneity is more frequent in meta-analysis of continuous than binary outcomes. *J Clin Epidemiol.* 2016;70:129–135.

17. Serghiou S, Goodman S. Random-effects meta-analysis: summarizing evidence with caveats. *JAMA.* 2019;321(3):301–302.

18. Presser S. The role of doubt in conceiving research: reflections from a career shaped by a dissertation. *Annu Rev Sociol.* 2022;48(1):1–21.

19. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet.* 1991;337(8746):867–872.

20. Hemkens L, Contopoulos-Ioannidis D, Ioannidis J. Routine collected data and comparative effectiveness evidence: promises and limitations. *CMAJ.* 2016;188(8):E158–E164.

21. Kaplan R, Chambers D, Glasgow R. Big data and large sample size: a cautionary note on the potential for bias. *Clin Transl Sci.* 2014;7(4):342–346.

22. Bradley S, DeVito N, Lloyd K, et al. Reducing bias and improving transparency in medical research: a critical overview of the problems, progress and suggested next steps. *J R Soc Med.* 2020;113(11):433–443.

23. Wang M, Bolland M, Grey A. Reporting of limitations of observational research. *Jama Intern Med.* 2015;175(9):1571–1572.

24. Kukull W, Ganguli M. Generalizability. *Neurology.* 2012;78(23):1886–1891.

25. Talari K, Goyal M. Retrospective studies – utility and caveats. *J R Coll Physicians Edinb.* 2020;50(4):398–402.

26. Tofthagen C. Threats to validity in retrospective studies. *J Adv Pract Oncol.* 2012;3(3):181–183.

27. Boston University Cohort studies. Available at: https://sphweb.bumc.bu.edu/otlt/mph-modules/ep/ep713_cohortstudies/index.html.

28. Röhrig B, Du Prel J, Wachtlin D, Kwiecien R, Blettner M. Small size calculation in clinical trials. *Dtsch Arztebl Int.* 2010;107(31–32):552–6. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2933537/pdf/Dtsch_Arztebl_Int-107-0552.pdf.

29. Vasileiou K, Barnet J, Thorpe S, Young T. Characterizing and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC Medical Research Methology.* 2018;18:148.

30. Button K, Ioannidis J, Mokrysz C, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci.* 2013;14(5):365–376.

31. Faber J, Fonseca L. How sample size influences research outcomes. *Dental Press J Orthod.* 2014;19(4):27–9.

32. Nosek B, Spies J, Motyl M. Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspect Psychol Sci.* 2012;7(6):615–631.

33. Jager KJ, Tripepi G, Chesnaye NC, Dekker FW, Zoccali C, Stel VS. Where to look for the most frequent biases? *Nephrology (Carlton).* 2020;25(6):435–441.

34. Lambert J. How to assess bias in clinical studies? *Clin Orthop Relat Res.* 2011;469(6):1794–1796.

35. Croskerry P. 50 Cognitive and affective biases in medicine. *Critical Thinking Program.* Dalhousie University: Halifax, Nova Scotia. Available at: https://sjrhem.ca/wp-content/uploads/2015/11/CriticaThinking-Listof50-biases.pdf.

36. Pannucci C, Wilkins E. Identifying and avoiding bias in research. *Plast Reconstr Surg.* 2010;126(2):619–625.

37. Heick T. The cognitive biases list: a visual of 180+ heuristics. Available at: https://www.teachthought.com/critical-thinking/cognitive-biases/.

38. Munafò M, Nosek B, Bishop D, et al. A manifesto for reproducible science. *Nat Hum Behav.* 2017;1:0021.

39. Murad M, Chu H, Lin L, Wang Z. The effect of publication bias magnitude and direction on the certainty in evidence. *BMJ Evid Based Med.* 2018;23(3):84–86.

40. Hammond M, Stehlik J, Drakos S, Kfoury A. Bias in medicine: lessons learned and mitigation strategies. *J Am Coll Cardiol Basic Trans Science.* 2021;6:78–85.

41. Doherty T, Carroll A. Believing in overcoming cognitive biases. *AMA J Ethics.* 2020;22(9):E773–778. Available at: https://journalofethics.ama-assn.org/sites/journalofethics.ama-assn.org/files/2020-08/medu1-2009.pdf

42. Tipton E, Hallberg K, Chan W. Implications of small samples for generalization adjustments and rules of thumb. Available at: https://journals.sagepub.com/doi/10.1177/0193841X16655665?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed or https://files.eric.ed.gov/fulltext/ED562348.pdf.

43. Alvarez G, Núñez-Cortés R, Solà I, et al. Sample size, study length, and inadequate controls were the most common self-acknowledged limitations in manual therapy trials: a methodological review. *J Clin Epidemiol.* 2021;130:96–106. Available at: https://doi.org/10.1016/j.jclinepi.2020.10.018 or https://www.comecollaboration.org/es/wp-content/uploads/sites/5/2020/11/1-s2.0-S0895435620311574-main-2.pdf.

44. Walker E, Hernandez AV, Kattan MW. Meta-analysis: its strengths and limitations. *Cleve Clin J Med.* 2008;75(6):431–439.

45. Epidemiology for the Uninitiated. Available at: https://www.bmj.com/about-bmj/resources-readers/publications/epidemiology-uninitiated.

46. Research Guides USC Libraries. Available at: https://libguides.usc.edu/writingguide/limitations#:∼:text=Definition,the%20findings%20from%20your%20research.

47. Fanelli D. Negative results are disappearing from disciplines and countries. *Scientometrics.* 2012;90(3):891–904. Available at: https://devaka.info/wp-content/uploads/2016/08/Fanelli12-NegativeResults.pdf.

48. Begg C, Berlin J. Publication bias: a problem in interpreting medical data. *Journal of the Royal Statistical Society Series A; Statistics in Society.* 1988;151(3):419–445.

49. Mlinaic A, Horvat M, Smolcic V. Dealing with the positive publication bias: why you should really publish your negative results. *Biochem Med (Zagreb).* 2017;27:447–452.

50. Nathan H, Pawlik TM. Limitations of claims and registry data in surgical oncology research. *Ann Surg Oncol.* 2008;15(2):415–423.

51. Bagaev E, Ali A, Saha S, et al. Hybrid surgery for severe mitral valve calcification: limitations and caveats for an open transcatheter approach. *Medicina.* 2022;58(1):93.

52. Ng T, McGory M, Ko C, Maggard M. Meta-analysis in surgery methods and limitations. *Arch Surg.* 2006;141(11):1125–30; discussion 1131.

53. ter Riet G, Chesley P, Gross AG, et al. All that glitters isn't gold: a survey on acknowledgment of limitations in biomedical studies. *PLoS One.* 2013;8(11):e73623.

54. Nunan D, Bankhead C, Aronson J. Selection bias. Catalogue of Bias. 2017. Available at: https://catalogofbias.org/biases/selection-bias/ and https://catalogofbias.org/.

55. Bornhöft G, Maxion-Bergemann S, Wolf U, et al. Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. *BMC Med Res Methodol.* 2006;6:56.

56. Gray J. Discussion and conclusion. *AME Med J.* 2019;4:26–26. Available at: https://amj.amegroups.org/article/view/4955/pdf.

57. Yavchitz A, Ravaud P, Hopewell S, Baron G, Boutron I. Impact of adding a limitations section to abstracts of systematic reviews on readers' interpretation: a randomized controlled trial. *BMC Med Res Methodol.* 2014;14:123.

58. Keserlioglu K, Kilicoglu H, ter Riet G. Impact of peer review on discussion of study limitations and strength of claims in randomized trial reports: a before and after study. *Res Integr Peer Rev.* 2019;4:19.

59. Morris A, Ioannidis J. Limitations of medical research and evidence at the patient-clinician encounter scale. *Chest.* 2013;143(4):1127–1135.

60. Lambert W, Janniger E. Limitations of randomized, controlled, double-blinded studies in determining safety and effectiveness of treatments. *Curr Med Res Opin.* 2022;38(6):1045–1046.

61. Mohseni M, Ameri H, Arab-Zozani M. Potential limitations in systematic review studies assessing the effect of the main intervention for treatment/therapy of COVID-19 patients: an overview. *Front Med (Lausanne)).* 2022;9:966632. Available at: https://www.frontiersin.org/articles/10.3389/fmed.2022.966632/full.

62. Hartman J, Forsen J, Wallace M, Neely G. Tutorial in clinical research: part IV: recognizing and controlling bias. *Laryngoscope.* 2002;112(1):23–31.

63. Delgado-Rodríguez M, Llorca J. Bias. *J Epidemiol Community Health.* 2004;58:635–41.

64. Emerson G, Warme W, Wolf F, Heckman J, Brand R, Leopold S. Testing for the presence of positive-outcome bias in peer review: a randomized controlled trial. *Arch Intern Med.* 2010;170(21):1934–1939.

65. International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. *N Engl J Med.* 1997;336(4):309–315.

66. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. Updated May 2023. Available at: https://www.icmje.org/icmje-recommendations.pdf.

67. Lingard L. The art of limitations. *Perspect Med Educ.* 2015;4(3):136–7.

68. Infante-Rivard C, Cusson A. Reflection on modern methods: selection bias – a review of recent developments. *Int J Epidemiol.* 2018;47(5):1714–1722.

69. Garrubba M, Joseph C, Melder A. Best practice to identify and prevent cognitive bias in clinical decision-making: scoping review. Centre for Clinical Effectiveness, Monash Innovation and Quality, Monash Health, Melbourne, Australia. Available at: https://dl.pezeshkamooz.com/pdf/cme/MedicalErrors/cognitive-bias-Dr.moradi.pdf.

70. McGauran N, Wieseler B, Kreis J, Schüler Y-B, Kölsch H, Kaiser T. Reporting bias in medical research - a narrative review. *Trials.* 2010;11(1):37. Available at: https://trialsjournal.biomedcentral.com/articles/10.1186/1745-6215-11-37?ref=data.gripe.

71. Morgado F, Meireles J, Neves C, Amaral A, Serreira M. Scale development: ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexao e*

Critica. 2017;30(3). Available at: https://www.scielo.br/j/prc/a/M6fHHwfVFM9GkjCrsr9SFHz/?lang=en.

72. Sheridan D, Julian D. Achievements and limitations of evidence-based medicine. *J Am Coll Cardiol.* 2016;68(2):204–213.

73. Feynman R. Cargo cult science. *Engineering and Science.* 1974;37:10–13. Available at: http://calteches.library.caltech.edu/51/2/CargoCult.pdf.

74. U.S. Department of Defense News Transcript. February 12 2002. Available at: https://archive.ph/20180320091111/http://archive.defense.gov/Transcripts/Transcript.aspx?TranscriptID=2636.

75. Resnik D, Shamoo A. Reproducibility and research integrity. *Account Res.* 2017;24(2):116–123. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5244822/pdf/nihms841390.pdf.

76. Hegedus E, Moody J. Clinimetrics corner: the many faces of selection bias. *J Man Manip Ther.* 2010;18(2):69–73.

77. Gelman A, Loken E. The statistical crisis in science. *Am Sci.* 2014;102(6):460–465. Available at: http://www.psychology.mcmaster.ca/bennett/psy710/readings/gelman-loken-2014.pdf.

78. Ott D. Hedging, weasel words, and truthiness and scientific writing. *JSLS.* 2018;22(4):e2018.00063.

79. Ott D. Words representing numeric probabilities in medical writing are ambiguous and misinterpreted. *JSLS.* 2021;25(3): e2021.00034.

80. U.S. Food & Drug Administration MAUDE. Manufacturer and user facility device experience. Available at: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfmaude/textsearch.cfm.

Appendix

How to calculate the Limitation Index Score (LIS). See **Tables 1–5**. Each limitation admitted to by authors in the article equals (=) one (1) point. Limitations may be generally stated by the author as a broad category, but can have multiple components, such as a retrospective study with these limitation components: 1. data or recall not accurate, 2. data missing, 3. selection bias not controlled, 4. confounders not controlled, 5. no randomization, 6. no blinding, 7. difficult to establish cause and effect, and 8. cannot draw a conclusion of causation. For each component, no matter how many are not explained and corrected, add an additional one (1) point to the score. See **Table 2**.

Each limitation not admitted to = two (2) points. A meta-analysis study gets an automatic 2 points since they are retrospective and have detrimental components that should be added to the 2 points. Data from insurance, state, national, Medicare, and Medicaid, because of incorrect coding, over-reporting, and underreporting, etc., score 2 points, and each component adds one additional point. Surveys and questionnaires get 2 points, and add one additional point for each bias. See **Table 3**.

Manufacturer and User Facility Device Experience (MAUDE) database articles receive an automatic five (5) points. The FDA access data site says, submissions can be "incomplete, inaccurate, untimely, unverified, or biased" and "the incidence or prevalence of an event cannot be determined from this reporting system alone due to underreporting of events, inaccuracies in reports, lack of verification that the device caused the reported event, and lack of information" and "MDR data alone cannot be used to establish rates of events, evaluate a change in event rates over time or compare event rates between devices. The number of reports cannot be interpreted or used in isolation to reach conclusions."[80] See **Table 4**. Add one additional point for each additional limitation noted in the article.

Add one additional point for each additional limitation and one point for each of its components. Extra blanks are for additional limitations and their component scores.

Table 1.
The Limitation Scoring Index is a Numeric Limitation Risk Assessment Score to Rank Risk Categories and Discounting Probability of Validity and Conclusions. The More Limitations in a Study, the Greater the Risk of Unreliable Findings and Conclusions

Number of limitations	Word description of discounting	Proposed percent discounting of conclusions	Outcome probability	Increasing level of less reliable conclusions
0	Unknown unknowns	1–10%	May have valid conclusion(s)	Warning
1–2	Some	15–25%	↓	↓
3–4	Probable	35–45%	↓	Caution
5–6	Likely	70–80%	↓	↓
7–8	Highly likely	85–95%	↓	↓
>8	Certain	97–100%	Very questionable conclusion(s)	Danger

Table 2.

Limitations May Be Generally Stated by the Author but Have Multiple Components, Such as a Retrospective Study Having Disadvantage Components of 1. Data or Recall Not Accurate, 2. Data Missing, 3. Selection Bias Not Controlled, 4. Confounders Not Controlled, 5. No Randomization, 6. No Blinding, 7 Difficult to Establish Cause and Effect, 8. Results Are Hypothesis Generating, and 9. Cannot Draw a Conclusion of Causation. For Each Component, Not Explained and Corrected, Add an Additional One (1) Point Is Added to the Score. Extra Blanks Are for Additional Limitations

---

One point for each limitation

One additional point for each component of each limitation

Retrospective study

Small sample size

Not generalizable

Selection bias

Not controlling for confounders

Not controlling for comorbidities

Incomplete or missing data

No long-term follow-up

Reporting errors

Measurement problems

Study design problems

Lack of standardized treatment

Subtotal for Table 2

Table 3.

An Automatic 2 Points is Added for Meta-Analysis Studies Since They Have All the Retrospective Detrimental Components.[44] Data from Insurance, State, National, Medicare, and Medicaid, Because of Incorrect Coding, Over Reporting, and Under-Reporting, Etc. Each Component of the Limitation Adds One Additional Point. For Surveys and Questionnaires Add One Additional Point for Each Bias. Extra Blanks Are for Additional Limitations

Two points for these limitations

One additional point for each limitation and one additional point for each limitation component.

Meta-analysis

Data from Medicare, Medicaid, insurance companies, disease, state, and national databases

Surveys and questionnaires

Each limitation not admitted to

Subtotal for Table 3

Table 4.

Automatic Five (5) Points for Manufacturer and User Facility Device Experience (MAUDE) Database Articles. The FDA Access Data Site Says Submissions Can Be "Incomplete, Inaccurate, Untimely, Unverified, or Biased" and "the Incidence or Prevalence of an Event Cannot Be Determined from This Reporting System Alone Due to Under-Reporting of Events, Inaccuracies in Reports, Lack of Verification That the Device Caused the Reported Event, and Lack of Information" and "DR Data Alone Cannot Be Used to Establish Rates of Events, Evaluate a Change in Event Rates over Time or Compare Event Rates between Devices. The Number of Reports Cannot Be Interpreted or Used in Isolation to Reach Conclusions"[80]

Five points for MAUDE based articles

One additional point for each additional limitation and one point for each of its components.

Subtotal for Table 4

Table 5.
Total Limitation Index Score

Limitations	Calculation
Subtotal for Table 2	
Subtotal for Table 3	
Subtotal for Table 4	
Total Limitation Index Score	