Clinical Nutrition 39 (2020) 1284–1291

Contents lists available at ScienceDirect

Clinical Nutrition

journal homepage: http://www.elsevier.com/locate/clnu

Original article

The fragility of statistically significant results from clinical nutrition randomized controlled trials

Michał Pędziwiatr [a, b, *, 1], Magdalena Mizera [a, 1], Michał Wysocki [a, b], Piotr Małczak [a, b], Tomasz Stefura [a], Natalia Gajewska [a], Grzegorz Torbicz [a], Jakub Droś [a], Artur Kacprzyk [a], Piotr Major [a, b], Stanisław Kłęk [c], Małgorzata Bała [d]

[a] 2nd Department of General Surgery, Jagiellonian University Medical College, Kraków, Poland
[b] Centre for Research, Training and Innovation in Surgery (CERTAIN Surgery), Kraków, Poland
[c] General and Oncology Surgery Unit, Stanley Dudrick's Memorial Hospital, Skawina, Poland
[d] Chair of Epidemiology and Preventive Medicine, Department of Hygiene and Dietetics, Jagiellonian University Medical College, Kraków, Poland

ARTICLE INFO

Article history:
Received 4 March 2019
Accepted 23 May 2019

Keywords:
Fragility index
Nutrition
Surgery
Randomized clinical trial
RCT

SUMMARY

Background & aims: Recently, a parameter called "Fragility index" (FI) has been proposed, which measures how many events the statistical significance relies on. The lower the FI the more "fragile" the results, and thus more care should be taken when interpreting the results. Our aim in this study was to check FI of nutritional trials.

Methods: We conducted a systematic review of human clinical nutrition RCTs that report statistically significant dichotomous primary outcomes. We searched the EMBASE, MEDLINE, and Scopus databases. The FI of primary outcomes using the Fisher exact test was calculated and checked the correlations of FI with the number of randomised trials, the p-value of primary outcomes, the publication date, the journal impact factor and the number of patients lost to follow-up.

Results: The initial database search revealed 5790 articles, 37 of which were included in qualitative synthesis. The median (IQR) FI for all studies was 1 (1–3). 28 studies (75.7%) had an FI lower or equal to 2, and in 12 (32.43%) articles, the FI was lower than the number of patients lost to follow-up. No correlations were found between FI and the study characteristics (number of randomized patients, p value of primary outcome, event ratio in experimental group, event ratio in control group, publication date, journal impact factor, lost to follow-up).

Conclusion: The results of RCTs in nutritional research often rely on a small number of events or patients. The number of patients lost to follow-up is frequently higher than the FI calculation. Formulating recommendations based on RCTs should be done with caution and FI may be used as auxiliary parameter when assessing the robustness of their findings.

1. Introduction

Although nutrition interventions have been studied for centuries, modern nutritional research is relatively young [1]. The first reports on the relationship between weight loss and surgical outcomes were published only 80 years ago [2]. Since that time various clinical nutrition formulas have been investigated and have shown to be beneficial in certain clinical situations. The evidence for the use of clinical nutrition as supportive to other treatments or as the primary intervention comes mostly from randomised controlled trials (RCTs) published within the last two decades.

It seems apparent that clinical nutrition studies work perfectly as RCTs – in the simplest design an experimental group is exposed to treatment and control groups are given a treatment or a placebo. For this reason, RCTs are regarded as the definitive study design for proving causality. In order to prove the effect of an intervention in a trial, statistical analyses are performed, and they must include p-values to determine the strength of observation against the null hypothesis or in other words whether an observation is a result of a change that was made by the intervention or is a result of random occurrences. However, these metrics are far from ideal and have

* Corresponding author. 2nd Department of General Surgery, Jagiellonian University Medical College, Kopernika 21, 31-501 Kraków, Poland. Fax: +48 12 421 34 56.
E-mail address: michal.pedziwiatr@uj.edu.pl (M. Pędziwiatr).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.clnu.2019.05.024

been recently criticised as too simplistic [3–8]. Therefore, other metrics are used (e.g. 95% confidence intervals) as auxiliary in determining the existence of a true treatment effect.

Recently, a new parameter called “Fragility Index” (FI) has been proposed as a valuable tool to establish the confidence in trial results. FI is used for dichotomous variables that show statistically significant difference between trial groups. It indicates the minimum number of patients whose status would have to change from a non-event to an event required to turn a statistically significant result into a non-significant result. The lower the FI, the more “fragile” (less statistically robust) the results of an RCT, and for this reason they must be interpreted more carefully. This approach has been already investigated in some previously published reviews [3,9].

There is a large number of RCTs in clinical nutrition and the majority show no or small differences in outcomes. What is more, international study groups formulate recommendations based on those trial results. Therefore, we hypothesised that trials reporting any changes rely on a small number of patients or events and are therefore characterised by low FI. Our primary aim in this study was to review RCTs in the field of clinical nutrition with statistically significant primary outcomes and to calculate the FI for those trials. Our secondary aim was to assess the relationship between trial characteristics and FI.

2. Materials and methods

We conducted a systematic review of human clinical nutrition RCTs following the methodology described in the Cochrane Handbook [10]. Reporting of our review follows the PRISMA Statement [11,12]. This systematic review was registered, and the detailed study protocol was published in the International Prospective Register of Ongoing Systematic Reviews (PROSPERO) under registration number CRD42018114712.

2.1. Search strategy

Our literature search included the EMBASE, MEDLINE, and Scopus databases. The search terms used were “random”, “randomised control trial”, “clinical trial”, “therapy diet”, “clinical nutrition”, “enteral nutrition”, “parenteral nutrition”, “immune nutrition”, “immunonutrition”, “pharmaconutrition”, “artificial feeding”, and combinations of these with Boolean “AND” and “OR” operators. We only searched for the articles published in English after 1 January 1992. Our most recent search was performed on 17 July, 2018. The full search strategy (OVID) and study flow chart presents as follows:

((random or randomized control*trial or clinical trial) not child*).ab,kw,sh,ti. AND ((((((((((((((((((((((((artificial adj3 feeding) or diet) adj3 therapy) or diet*) adj3 supplement*) adj3 immune nutrition) or enter*) adjimmunonutrition) or enter*) adj3 Nutrition) or enter*) adj3 feeding) or IMMUN* NUTRITION or immune?enhanc* enteral) adj3 nutrition) or immunoenhanced enteral) adj3 nutrition) or Immunomodulation* or immunonutrition or nutrition*) ajd3 support) or oral) adj3 supplement*) or Parenteral) adj3 Nutrition) or perioperative) adj3 nutrition) or PHARMACONUTRITION).ab,kw,sh,ti.

2.2. Selection of studies

To select relevant RCTs, PICO eligibility criteria were used. We included all randomised control trials involving human patients undergoing nutritional interventions (in comparison to other interventions or placebo) and reporting statistically significant dichotomous primary outcomes. The outcome was considered primary when it was specified as primary the author of an article. If this information was not provided, an outcome for which study sample calculation was performed was considered as primary outcome for the study. If neither of those was present, the first statistically significant dichotomous outcome was considered ‘primary’. If primary outcome of a study was composite and only one of its components was significant whereas the entire composite outcome was not statistically significant, then we decided to exclude such study. We also excluded all studies that used a quasi-randomised or non-randomised methodology or if the article was not in English or if the study was not full-length article (e.g. conference abstract, poster etc.).

We defined nutritional intervention as either nutritional treatment (the quantity or quality of calories delivered or the route of administration) or immune nutrition (for example polyunsaturated fatty acids or immune enhancing amino acids). Research on dietary supplements was not included. In order to allow an FI calculation, the variable had to be dichotomous, which means it had to take one of only two possibilities (e.g. mortality, morbidity, occurrence of adverse events, etc.). The variable was statistically significant when the p-value was 0.05 or less.

After calibration with pilot-tested electronic forms, screening was performed with Covidence online software. The eligibility and data extraction were assessed by three pairs of reviewers (JD, NG, AK, MM, TS, GT).The reviewers received training and detailed oral and written instruction on how to perform title and abstract and full-text screening, including meetings with calibration and sample data review exercises. Differences in opinion were resolved by discussion until a consensus was reached.

2.3. Data extraction and measured outcomes

The data were extracted on first author, year of publication, title, name of publishing journal and its impact factor in the year of publication according to Journal Citation Reports by Clarivate. None of the articles included was published in 2018. We also extracted the following information: trial sample size (number of randomised patients) and number of participating centres, significance of primary outcome (Yes/No), p-value for primary outcome, study sample calculation (yes/no), and funding source (medical consortium/the government/no funding/not stated). The number of events (event ratio) and total number of patients in each experimental and control group were recorded. If the article reported both an intention to treat (ITT) and per protocol (PP) analysis, we used the ITT data for the FI calculation. Trials were assigned into one of the following categories: nutritional intervention (enteral/parenteral/immunonutrition) vs. control (enteral/parenteral/immunonutrition/placebo) group. Patients were classified as intensive care/surgical/pancreatitis/non-surgical depending on their underlying disease and primary treatment. Patients lost to follow-up were noted if reported. Details of included studies are provided in Supplementary material 1.

2.4. Fragility index calculation

The FI was calculated as described by Walsh et al. [3]. Statistically significant results of the trials were extracted and presented in a two-by-two contingency table, consisting of the number of events and total number of patients in experimental (receiving intervention) and control group. An event was added to the group with the smaller number of events (and a non-event from the same group was subtracted in order not to change the patients group number). The two-sided p-value for the Fisher's exact test was then recalculated. This procedure was reiterated until the p-value became

equal to or greater than 0.05. The FI is indicated by the number of events required to change the significance of the result.

2.5. Quality assessment

We used the Cochrane risk of bias tool (RoB) to assess the risk of bias of the RCTs included. The tool addresses specific domains that are assessed as “high risk of bias”, “low risk of bias”, or “unclear”. There are seven main domains: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reports, and other bias. Depending on the risk of bias assessment for specific domains overall, the risk of bias of each study was classified as low if all criteria were met (i.e. low risk of bias for each domain) or one criterion was unclear. Alternatively, studies were classified as high risk of bias if one criterion was not met (i.e. high risk of bias for one domain) or two or more criteria were unclear. The risk of bias assessment across studies is provided in Supplementary material 2.

2.6. Statistical analysis

All analyses were conducted using Statsoft STATISTICA 13.0 (Statsoft Inc., Tulsa, Oklahoma, USA). Continuous variables are reported as means with standard deviations or medians with inter quartile ranges (IQR) depending on consistency with normal distribution (based on Student's t-test and Mann–Whitney's test when appropriate). The Spearman's correlation and linear regression models were also used. Categorical variables are presented as counts and percentages. To compare categorical data, Pearson's Chisquared, Yates' and Fisher's tests were used when appropriate. Statistical significance was observed at $p < 0.05$.

3. Results

3.1. Search results

The initial database search revealed 5790 articles. Title and abstract screening yielded 469 full-text articles for the assessment. Finally, 37 articles were included in qualitative synthesis (Supplement 1) [13–49]. Details of the systematic review and data extraction are provided in the PRISMA Flowchart (Fig. 1).

3.2. Fragility index

The statistically significant primary outcome was present in all studies. Median FI for all studies was 1 (IQR1–2). The FI distribution is presented in Fig. 2.

28 studies (75.7%) had an FI lower or equal to 2. All studies reported those lost to follow-up. Out of all the studies included in 12 (32.43%) articles, the FI was lower than the number of patients lost to follow-up.

3.3. Trial characteristics

The median trial size was 100.0 patients (IQR 50.0–159.0).The median number of centres participating in the study was 10 (IQR 6–17). The median number of events was 12.0 (IQR 5.0–24.0) in the experimental group and 18.0 (IQR 12.0–33.0) in the control group. 8 (21.6%) trials were multicentre, 27 (73.0%) were single centre, and 2 (5.4%) trials did not provide data about the single/multicentre design. 22 (48.9%) trials were blinded; 15 (40.5%) trials were sponsored by a medical consortium, 18 (48.7%) declared no funding, and 4 (10.8%) did not provide data on type of funding.

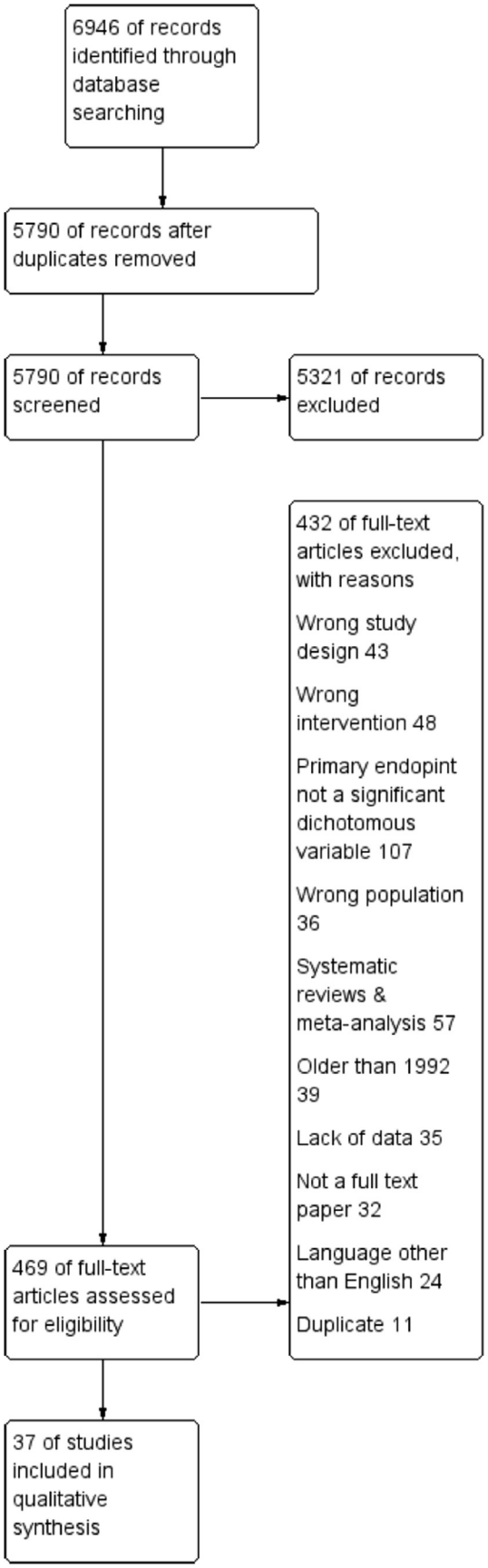

Fig. 1. Study flowchart.

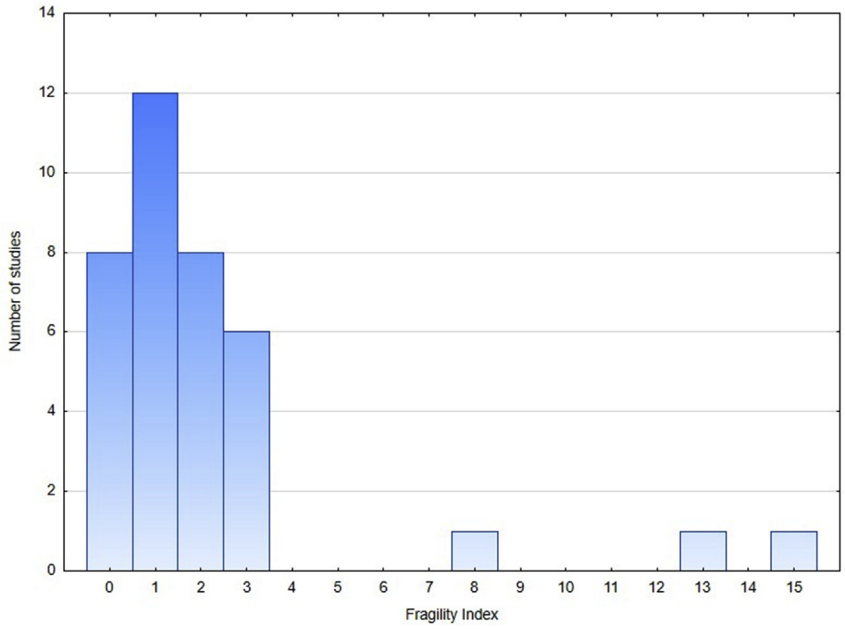

Fig. 2. Fragility index distribution for primary outcomes.

3.4. Fragility index association with selected variables

Larger trials (equal or greater than the median, 100 or more patients randomised) did not have significantly greater FI than smaller trials (less than median, <100 randomised patients) ($p = 0.975$).Trials with a p-value of 0.01–0.05 had a median FI of 1 (IQR 0–2) and those with a p-value 0.01–0.001 had a median FI of 2.5 (IQR 1.5–3); the difference was statistically significant ($p = 0.047$). We also did not find any differences in FI between studies published within the last 10 years or those published earlier than 10 years ago ($p = 0.796$), or any differences between trials sponsored by medical company vs. other source of funding ($p = 0.751$). We also did not find any statistical differences in FI between single and multicentre trials ($p = 0.441$) or when the study sample size was calculated or not ($p = 0.094$).When calculation FI scores for one type of outcome (mortality, morbidity, infections and others) they were: not statistically different.

3.5. Quality assessment

Quality assessment results are presented in Fig. 3 (Supplement 2). 11 studies were classified as having an overall low risk of bias (29.7%). Studies classified as having an overall low risk of bias had an FI of 3 (IQR 1–3), while studies classified as having an overall high risk of bias had an FI of 1 (IQR0–2); this difference was significant ($p = 0.034$).

3.6. Fragility index correlations with study characteristics

In Spearman's ranks, no correlations were found between FI and the number of randomized patients, p value of primary outcome, event ratio in experimental group, event ratio in control group, publication date, journal impact factor, lost to follow-up. The tested correlations of FI are presented in Table 1.

4. Discussion

In this review, we found that FI remains lower than or equal to 2 in three quarters of the trials included. This simply means that in the majority of RCTs on clinical nutrition the results are fragile, i.e. only two events are sufficient to change the significance of the trial findings and its conclusion. In addition, in one third of the studies the FI was lower than the number of patients lost to follow-up, indicating a potential problem of altering the final results by

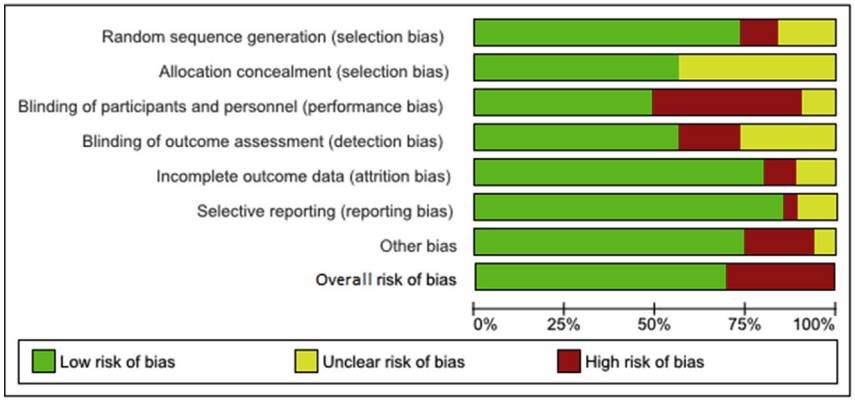

Fig. 3. Quality assessment of included studies.

Table 1
Fragility index Spearman's correlation with selected variables.

	R	p-value
Number of randomized patients	0.165	0.330
p value of primary outcome	−0.191	0.257
Event ratio in experimental group	−0.015	0.930
Event ratio in control group	0.187	0.269
Publication date	−0.119	0.482
Journal impact factor	0.124	0.498
Lost to follow-up	−0.147	0.384

including patients lost to follow-up, which has also been shown previously [3,9]. Therefore, we confirmed that even though the results of RCTs in clinical nutrition show a statistically significant effect of a given intervention for primary outcomes, which is confirmed by the p-value, those results usually depend on a small number of patients.

In order to unambiguously present study conclusions, the quality of published trials continues to improve. Although p-values are the gold standard in the presentation of results, they have been criticised as too simplistic and are usually accompanied by 95% confidence intervals [3]. These not only allow the reader to assess the significant difference between studied groups but also the magnitude of the treatment effect [50]. Moreover, just because the treatment effect is significantly larger in one group, it does not necessarily mean that this difference is clinically meaningful. This is particularly important in clinical nutrition, which in the majority of clinical scenarios is used as auxiliary rather than primary therapy.

Our findings are important for several reasons. FI is probably the only parameter that facilitates a clinician's appreciation of the robustness of findings by providing them with the exact number of patients required to change the significance of the result. Studies with greater FI are considered less fragile: their results are more solid and resistant to changes or loss to follow-up. Although FI represents the strength of the result in the numeric sense, it is not necessarily associated with the clinical relevance of the result. Moreover, we did not find any correlation between FI and study characteristics. This is not in line with other reviews where a sample size was correlated with the total number of events, sample size [51–53]. On the other hand, it is important to remember that the study methodology and the journal impact factor, as well as the source of funding, has only a limited influence on quality of RCTs as shown in a review by Ahmed Ali et al. [54]. In addition, the presentation of FI with no relation to study sample and number of events may be also misleading. For example, an FI of 4 for a sample size of 20, as compared to a sample size of 200, shows that the clinical relevance depends strongly on the size of the trial as well as number of events. The FI score divided by the total study sample size is called the Fragility Quotient and is a derivative parameter that may be also used in the assessment of results.

Although in RCTs with statistically significant dichotomous outcomes the FI is usually greater than zero, it is mathematically possible to achieve an FI of 0. To calculate this parameter, we followed the methodology by Walsh et al. [3] who originally used the Fisher's exact test which is useful for categorical data that result from classifying objects in two different ways. On the contrary to other statistical methods for this type of variables (such as a Chi squared Pearson's or Chi-squared Walsh's tests being the of logistic regression models) Fisher's test is considered the most conservative and seems to be the most appropriate for comparisons of dichotomous variables (although it is also is more prone to type II error).As a result, it can also lead to a situation when FI = 0 despite statistically significant difference between groups in the original study and clearly confirms the lack of robustness of its results. Obviously, a question arises whether FI scores would be different when tests other than Fisher's were used for their calculations. It is an important and interesting topic that should be evaluated in the future. However, it was beyond the topic of the current review.

Although this study is the first to investigate FI in the field of clinical nutrition, previously there have been similar reports published in other disciplines of medicine. Walsh et al. was probably the first to calculate FI in a larger number of RCTs that were published in five leading medical journals [3]. In his review, the FI was 3 or less in a quarter of the studies and loss to follow-up exceeded the FI in more than half the trials. The authors concluded that the statistically significant results of many RCTs hinge on small numbers of events. Therefore, FI complements the p-value and helps identify less robust results. Ridgeon et al. showed even more unexpected calculations from RCTs in critical care, reporting statistically significant effects on mortality [9]. More than 40% of trials had an FI of less than or equal to 1. Other reports are in line with these findings [52,55–57]. A study of emergency medicine, urology, paediatric medicine and otolaryngology reported a median FI between 3 and 7. It seems that regardless of the medical discipline, the results of RCTs hinge on a small number of events and therefore reporting FI would help clinicians make better-informed decisions about the robustness of findings in these trials. Perhaps even more important is the aspect of patients lost to follow-up. In our review, one third of the studies had an FI smaller than the number of patients lost to follow-up, indicating that the outcomes for lost participants could have changed the results from significant to non-significant. According to the LOST-IT review, up to one-third of trials reporting significant results for dichotomous primary outcomes that are patient-important lose significance if making plausible assumptions about their loss to follow-up [58]. This, together with a relatively high percentage of trials reporting no such information, raises the question how we should deal with patients lost to follow-up. In our opinion, FI may help to establish whether the number of patients lost to follow-up jeopardises the reliability of conclusions and may further influence treatment recommendations.

The majority of RCTs included in our review were previously used to formulate guidelines, society recommendations or position papers. However, sometimes an RCT is used to formulate strong recommendations despite a low FI. Consider an example. In an RCT by Kłęk et al. (NCT01025414), 78 malnourished patients undergoing pancreatic resection who developed pancreatic fistula were randomised to groups receiving either enteral nutrition or total parenteral nutrition [59] for 30 days. Their primary outcome was 30-day pancreatic fistula closure rate. The rates in patients receiving enteral and parenteral nutrition were 60% (24 of 40) and 37% (14 of 38), respectively ($p = 0.043$). Calculations revealed that the FI of the trial was 1, meaning that adding one event to the parenteral treatment arm eliminated statistical significance. Moreover, a study sample calculation performed by the authors before the beginning of the trial indicated that at least 39 patients should be allocated to each of the two compared arms. However, the number of patients in the experimental group was 38. Therefore, we may consider the results of the study very fragile. What is more, in a position paper of the International Study Group on Pancreatic Surgery published in 2018, the results of this study were used to formulate a recommendation that nasojejunal feeding should be preferred over parenteral nutrition, because enteral feeding is associated with significantly greater closure rates and a shorter time to closure of pancreatic fistula [60].It is important to note that a situation where recommendations are based on a single, small sample trial with low FI score is nowadays relatively uncommon, however in such cases caution is required before considering change in clinical practice on the basis of such evidence [57]. Besides, although FI score seems

very convenient to guide clinicians about the robustness of RCTs' results, they cannot rely solely on this item. For instance, sample size and its calculation before the trial is initiated are the main parameters that need to be taken into account when analysing the robustness of evidence and anticipating type I or type II errors. Moreover, in certain clinical situation reaching required study sample may be very difficult or even unrealistic. It does not mean there is no place for underpowered trials as long as their design is appropriate. They still may be used in subsequent meta-analyses to increase the power and precision of the estimated intervention effects [10]. Moreover, in the planning of an RCT systematic reviews of existing evidence can be utilised and the results of a single trial should also be reported in the context of the whole body of evidence summarised by systematic review [61–65].

Interestingly, the majority of published meta-analyses include too few randomised patients, to reach appropriate statistical power leading to overestimating (type I) or underestimation (type II) of intervention effects [66,67]. Therefore, trial sequential analysis methodology has been proposed to handle the issue of the reliability of cumulative evidence derived from multiple, heterogeneous, often "fragile" studies. Briefly, trial sequential analysis calculates the required number of patients, based on the predefined anticipated intervention effect [68]. In other words, allows for meta-analytic interpretation of data, with better control of type I and type II errors than the traditional meta-analysis using confidence intervals. This is also another parameter that may be used in evaluation of available evidence.

There are several limitations to this study. First, in our search strategy we did not include studies on single substances (vitamins, micronutrients). We decided to include papers published in English only within the last 25 years. For this reason, we set a time frame starting from 1st January 1992). Our initial literature search was performed in 2017. It was later updated in July 2018 due to slower work progress than anticipated. We also set language limitations, therefore it is likely that we did not include a small proportion of relevant studies. Although we were able to calculate the FI in each trial, we did not assess the how clinically meaningful were the effects observed in the trial. In other words, low FI should not be confused with low clinical effect of treatment measured. Finally, due to the nature of FI we included only studies with dichotomous primary outcomes in 1:1 RCTs, which limits the applicability of this parameter. We deliberately refrained from FI analysis in secondary outcomes as, in our understanding, trial design aims to confirm differences between groups in primary outcomes rather than in secondary endpoints. In addition, we included studies in which authors defined more than one primary outcome, however, and in some of them study sample calculations were not performed for each of these primary outcomes. Moreover, according to Walsh et al. and Tignanelli et al. applying FI to time-to-event data in which the numbers of events in both groups are similar but there is a clear difference in the timing of the events may be inappropriate and result in finding such trials inappropriately fragile [3,57]. However, in all studies we included there was no trial in which the number of events was similar with the only difference being time-to-event. Finally, we have made an attempt to compare the difference specified by the authors in sample size calculations (absolute or relative) with the differences obtained on the basis of study results (risk difference or relative risk reduction). Unfortunately, in the majority of papers this information was imprecisely reported (e.g. not specified if the authors referred to absolute or relative difference, not enough data to allow calculations). In 21 studies we compared planned and obtained risk differences and relative risk reductions (Supplementary material 1). Only in 7 studies obtained values were equal or higher than estimated, whereas in the remaining they were lower. We deliberately resigned from analysing correlations between these values and FI scores since the sample group would be too small to draw any reliable conclusions.

We conclude that statistically significant results of dichotomous primary outcomes in majority of clinical nutrition randomised trials are fragile and usually depend on very small number of patients. Taking this into account and the fact that the FI is often smaller than loss to follow-up, we postulate that this parameter should be reported in well-designed RCTs. Together with traditional and commonly used trial characteristics (p-value, study sample, confidence intervals) as well as putting the findings in the context of the whole body of evidence might help clinicians make better-informed decisions about the robustness of findings reported by these trials and their implications for changes in clinical practice.

Author contributions

Michał Pędziwiatr: Conceptualization, Investigation, Methodology, Project administration, Supervision Roles, Writing – original draft. **Magdalena Mizera:** Data curation, Formal analysis, Investigation, Project administration, Writing – original draft. **Michał Wysocki:** Data curation, Formal analysis, Methodology, Software, Supervision, Visualization, Writing – review & editing. **Piotr Małczak:** Data curation, Formal analysis, Methodology, Software, Supervision, Visualization, Writing – review & editing. **Tomasz Sefura:** Data curation, Formal analysis, Investigation, Writing – review & editing. **Natalia Gajewska**: Data curation, Formal analysis, Investigation, Writing – review & editing. **Grzegorz Torbicz**: Data curation, Formal analysis, Investigation, Writing – review & editing. **Jakub Droś**: Data curation, Formal analysis, Investigation, Writing – review & editing. **Artur Kacprzyk**: Data curation, Formal analysis, Investigation, Writing – review & editing. **Piotr Major**: Conceptualization, Data curation, Investigation, Methodology, Writing – review & editing. **Stanisław Kłęk:** Conceptualization, Data curation, Investigation, Methodology, Writing – original draft. **Małgorzata Bała:** Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Supervision.

Conflict of interest

The authors report no conflicts of interest in this work.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clnu.2019.05.024.

References

[1] Mozaffarian D, Rosenberg I, Uauy R. History of modern nutrition science-implications for current research, dietary guidelines, and food policy. BMJ 2018;361:k2392. https://doi.org/10.1136/bmj.k2392. Jun.
[2] Studley H. Percentage of weight loss: a basic indicator of surgical risk in patients with chronic peptic ulcer. 1936. Nutr Hosp 2001;16:141–3. Jul-Aug.
[3] Walsh M, Srinathan SK, McAuley DF, Mrkobrada M, Levine O, Ribic C, et al. The statistical significance of randomized controlled trial results is frequently fragile: a case for a Fragility index. J Clin Epidemiol 2014;67:622–8. https://doi.org/10.1016/j.jclinepi.2013.10.019.
[4] Sterne J, Davey Smith G. Sifting the evidence-what's wrong with significance tests? BMJ 2001;322:226–31.
[5] Rozeboom W. The fallacy of the null-hypothesis significance test. Psychol Bull 1960;57:416–28.
[6] Rothman K. Significance questing. Ann Intern Med 1986;105:445–7.

[7] Goodman S. Toward evidence-based medical statistics. 1: the P value fallacy. Ann Intern Med 1999;130. 995–100.
[8] McIlroy D. Seduced by a P-value. Anaesth Intensive Care 2014;42:551–4.
[9] Ridgeon E, Young P, Bellomo R, Mucchetti M, Lembo R, Landoni G. The fragility index in multicenter randomized controlled critical care trials. Crit Care Med 2016;44:1278–84. https://doi.org/10.1097/CCM.0000000000001670. Jul.
[10] Higgins J, Green S. Collaboration. In: TC, editor. Cochrane Handbook for systematic reviews of interventions; 2011. Novemb. 25, 2015.
[11] Moher D, Liberati A, Tetzlaff J, Altman DG, Altman D, Antes G, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6. https://doi.org/10.1371/journal.pmed.1000097.
[12] Liberati A, Altman D, Tetzlaff J. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009:b2700. https://doi.org/10.1136/bmj.b2700.
[13] Brennan M, Pisters P, Posner M, Quesada O, Shike M. A prospective randomized trial of total parenteral nutrition after major pancreatic resection for malignancy. Ann Surg 1994;220:436–41. Oct.
[14] Fan S, Lo C, Lai E, Chu K, Liu C, Wong J. Perioperative nutritional support in patients undergoing hepatectomy for hepatocellular carcinoma. N Engl J Med 1994;(331):1547–52. Dec 8.
[15] Déchelotte P, Hasselmann M, Cynober L, Allaouchiche B, Coëffier M, Hecketsweiler B, et al. L-Alanyl-L-glutamine dipeptide-supplemented total parenteral nutrition reduces infectious complications and glucose intolerance in critically ill patients: the French controlled, randomized, double-blind, multicenter study. Crit Care Med 2006;34:598–604. https://doi.org/10.1097/01.CCM.0000201004.30750.D1.
[16] Eneroth M, Olsson UB, Thorngren KG. Nutritional supplementation decreases hip fracture-related complications. Clin Orthop Relat Res 2006:212–7. https://doi.org/10.1097/01.blo.0000224054.86625.06.
[17] Petrov MS, Kukosh MV, Emelyanov NV. A randomized controlled trial of enteral versus parenteral feeding in patients with predicted severe acute pancreatitis shows a significant reduction in mortality and in infected pancreatic complications with total enteral nutrition. Dig Surg 2006;23:336–45. https://doi.org/10.1159/000097949.
[18] Oguz M, Kerem M, Bedirli A, Mentes BB, Sakrak O, Salman B, et al. L-Alanin-L-glutamine supplementation improves the outcome after colorectal surgery for cancer. Color Dis 2007;9:515–20. https://doi.org/10.1111/j.1463-1318.2006.01174.x.
[19] Fuentes-Orozco C, Cervantes-Guevara G, Muciño-Hernández I, López-Ortega A, Ambriz-González G, Gutiérrez-De-La-Rosa JL, et al. L-Alanyl-L-glutamine-supplemented parenteral nutrition decreases infectious morbidity rate in patients with severe acute pancreatitis. J Parenter Enter Nutr 2008;32:403–11. https://doi.org/10.1177/0148607108319797.
[20] Badía-Tahull M, Llop-Talaverón J, Leiva-Badosa E, Biondo S, Farran-Teixidó L, Ramón-Torrell J, et al. A randomised study on the clinical progress of high-risk elective major gastrointestinal surgery patients treated with olive oil-based parenteral nutrition with or without a fish oil supplement. Br J Nutr 2010;104:737–41. https://doi.org/10.1017/S0007114510001066. Sep.
[21] Jiang Z, Wilmore D, Wang X, Wei J, Zhang Z, Gu Z, et al. Randomized clinical trial of intravenous soybean oil alone versus soybean oil plus fish oil emulsion after gastrointestinal cancer surgery. Br J Surg 2010;97:804–9. https://doi.org/10.1002/bjs.6999. Jun.
[22] Klek S, Sierzega M, Turczynowski L, Szybinski P, Szczepanek K, Kulig J. Enteral and parenteral nutrition in the conservative treatment of pancreatic fistula: a randomized clinical trial. Gastroenterology 2011;141:157–63. https://doi.org/10.1053/j.gastro.2011.03.040.
[23] Klek S, Sierzega M, Szybinski P, Szczepanek K, Scislo L, Walewska E, et al. The immunomodulating enteral nutrition in malnourished surgical patients – a prospective, randomized, double-blind clinical trial. Clin Nutr 2011;30:282–8. https://doi.org/10.1016/j.clnu.2010.10.001.
[24] Miyata H, Yano M, Yasuda T, Hamano R, Yamasaki M, Hou E. Randomized study of clinical effect of enteral nutrition support during neoadjuvant chemotherapy on chemotherapy-related toxicity in patients with esophageal cancer. Clin Nutr 2011;31:330–6. https://doi.org/10.1016/j.clnu.2011.11.002.
[25] Griffiths R, Jones C, Palmer T. Six-month outcome of critically ill patients given glutamine-supplemented parenteral nutrition. Nutrition 1997;13:295–302. Apr.
[26] Roth B, Birkhäuser FD, Zehnder P, Thalmann GN, Huwyler M, Burkhard FC, et al. Parenteral nutrition does not improve postoperative recovery from radical cystectomy: results of a prospective randomised trial. Eur Urol 2012;63:475–82. https://doi.org/10.1016/j.eururo.2012.05.052.
[27] Zhu M, Tang D, Hou J, Wei J, Hua B, Sun J, et al. Impact of fish oil enriched total parenteral nutrition on elderly 2012;125:178–81.
[28] Heidegger CP, Berger MM, Graf S, Zingg W, Darmon P, Costanza MC, et al. Optimisation of energy provision with supplemental parenteral nutrition in critically ill patients: a randomised controlled clinical trial. Lancet 2013;381:385–93. https://doi.org/10.1016/S0140-6736(12)61351-8.
[29] Aida T, Furukawa K, Suzuki D, Shimizu H, Yoshidome H, Ohtsuka M, et al. Preoperative immunonutrition decreases postoperative complications by modulating prostaglandin E2 production and T-cell differentiation in patients undergoing pancreatoduodenectomy. Surg 2014;155:124–33. https://doi.org/10.1016/j.surg.2013.05.040.
[30] Anbar R, Beloosesky Y, Cohen J, Madar Z, Weiss A, Theilla M, et al. Tight calorie control in geriatric patients following hip fracture decreases complications: a randomized, controlled study. Clin Nutr 2014;33:23–8. https://doi.org/10.1016/j.clnu.2013.03.005.
[31] Owais AE, Kabir SI, Mcnaught C, Gatt M, MacFie J. A single-blinded randomised clinical trial of permissive underfeeding in patients requiring parenteral nutrition. Clin Nutr 2014;33:997–1001. https://doi.org/10.1016/j.clnu.2014.01.005.
[32] Vidal-Casariego A, Calleja-Fernández A, De Urbina-González JJO, Cano-Rodríguez I, Cordido F, Ballesteros-Pomar MD. Efficacy of glutamine in the prevention of acute radiation enteritis: a randomized controlled trial. J Parenter Enter Nutr 2014;38:205–13. https://doi.org/10.1177/0148607113478191.
[33] Grau-Carmona T, Bonet-Saris A, García-De-Lorenzo A, Sánchez-Alvarez C, Rodríguez-Pozo A, Acosta-Escribano J, et al. Influence of n-3 polyunsaturated fatty acids enriched lipid emulsions on nosocomial infections and clinical outcomes in critically Ill patients: ICU lipids study. Crit Care Med 2015;43:31–9. https://doi.org/10.1097/CCM.0000000000000612.
[34] Li CH, Chen DP, Yang J. Enteral nutritional support in patients with head injuries after craniocerebral surgery. Turk Neurosurg 2015;25:873–6. https://doi.org/10.5137/1019-5149.JTN.9503-13.1.
[35] Moya P, Soriano-Irigaray L, Ramirez JM, Garcea A, Blasco O, Blanco FJ, et al. Perioperative standard oral nutrition supplements versus immunonutrition in patients undergoing colorectal resection in an enhanced recovery (ERAS) protocol. Med 2016;95:1–11. https://doi.org/10.1097/MD.0000000000003704.
[36] Kalfarentzos F, Kehagias J, Mead N, Kokkinis K, Gogos C. Enteral nutrition is superior to parenteral nutrition in severe acute pancreatitis: results of a randomized prospective trial. Br J Surg 1997;84:1665–9. Dec.
[37] Moya P, Miranda E, Soriano-Irigaray L, Arroyo A, Aguilar M del M, Bellón M, et al. Perioperative immunonutrition in normo-nourished patients undergoing laparoscopic colorectal resection. Surg Endosc 2016;30:4946–53. https://doi.org/10.1007/s00464-016-4836-7.
[38] Perinel J, Mariette C, Dousset B, Sielezneff I, Gainant A, Mabrut J-Y, et al. Early enteral versus total parenteral nutrition in patients undergoing pancreaticoduodenectomy: a randomized multicenter controlled trial (Nutri-DPC). Ann Surg 2016;264:731–7. https://doi.org/10.1097/SLA.0000000000001896.
[39] Petros S, Horbach M, Seidel F, Weidhase L. Hypocaloric vs normocaloric nutrition in critically ill patients. J Parenter Enter Nutr 2016;40:242–9. https://doi.org/10.1177/0148607114528980.
[40] Chen H, Wang W, Hong C, Zhang M, Hong Y, Wang S, et al. Omega-3 fish oil reduces mortality due to severe sepsis with acute gastrointestinal injury grade III. Pharmacogn Mag 2017;13:407–12. Jul-Sep.
[41] Chen H, Wang W, Hong Y, Zhang H, Hong C, Liu X. Single-blinded, randomized, and controlled clinical trial evaluating the effects of omega-3 fatty acids among septic patients with intestinal dysfunction: a pilot study. Exp Ther Med 2017;14:1505–11. Aug.
[42] Vijey Aanandhi M, Meha J. Enteral/oral glutamine supplementation in patients following surgery and accidental injury. Asian J Pharm Clin Res 2017;10:477–9. https://doi.org/10.22159/ajpcr.2017.v10i3.16569.
[43] Zhang B, Wei G, Li R, Wang Y, Yu J, Wang R, et al. n-3 fatty acid-based parenteral nutrition improves postoperative recovery for cirrhotic patients with liver cancer: a randomized controlled clinical trial. Clin Nutr 2017;36:1239–44. https://doi.org/10.1016/j.clnu.2016.08.002.
[44] Braga M. Perioperative immunonutrition in patients undergoing cancer surgery. Arch Surg 1999;134:428. https://doi.org/10.1001/archsurg.134.4.428.
[45] Bozzetti F, Gavazzi C, Miceli R, Rossi N, Mariani L, Cozzaglio L, et al. Perioperative total parenteral nutrition in malnourished, gastrointestinal cancer patients: a randomized, clinical trial. J Parenter Enter Nutr 2000;24:7–14.
[46] Bozzetti F, Braga M, Gianotti L, Gavazzi C, Mariani L. Postoperative enteral versus parenteral nutrition in malnourished patients with gastrointestinal cancer: a randomised multicentre trial. Lancet 2001;358:1487–92. https://doi.org/10.1016/S0140-6736(01)06578-3.
[47] Fuentes-Orozco C, Anaya-Prado R, González-Ojeda A, Arenas-Márquez H, Cabrera-Pivaral C, Cervantes-Guevara G, et al. L-Alanyl-L-glutamine-supplemented parenteral nutrition improves infectious morbidity in secondary peritonitis. Clin Nutr 2004;23:13–21. https://doi.org/10.1016/S0261-5614(03)00055-4.
[48] Ahrens CL, Barletta JF, Kanji S, Tyburski JG, Wilson RF, Janisse JJ, et al. Effect of low-calorie parenteral nutrition on the incidence and severity of hyperglycemia in surgical patients: A randomized, controlled trial. Crit Care Med 2005;33:2507–12. https://doi.org/10.1097/01.CCM.0000186746.64572.8A.
[49] Muto Y, Sato S, Watanabe A, Moriwaki H, Suzuki K, Kato A, et al. Effects of oral branched-chain amino acid granules on event-free survival in patients with liver cirrhosis. Clin Gastroenterol Hepatol 2005;3:705–13. https://doi.org/10.1016/S1542-3565(05)00017-0.
[50] Guyatt G, Rennie D, Meade MO, Cook DJ. Users' guides to the medical literature: a manual for evidence-based clinical practice, 3rd ed. AMA Press.
[51] Khormaee S, Choe J, Ruzbarsky J, Agarwal K, Blanco J, Doyle S, et al. The fragility of statistically significant results in pediatric orthopaedic randomized controlled trials as quantified by the fragility index: a systematic review. J Pediatr Orthop 2018;38:e418–23. https://doi.org/10.1097/BPO.0000000000001201. Sep.
[52] Skinner M, Tritz D, Farahani C, Ross A, Hamilton T, Vassar M. The fragility of statistically significant results in otolaryngology randomized trials. Am J Otolaryngol 2019:61–6. https://doi.org/10.1016/j.amjoto.2018.10.011. Jan-Feb.
[53] Narayan V, Gandhi S, Chrouser K, Evaniew N, Dahm P. The fragility of statistically significant findings from randomised controlled trials in the urological literature. BJU Int 2018;122:160–6. https://doi.org/10.1111/bju.14210. Jul.

[54] Ahmed Ali U, Reiber B, Ten Hove J, van der Sluis P, Gooszen H, Boermeester M, et al. Journal impact factor and methodological quality of surgical randomized controlled trials: an empirical study. Langenbecks Arch Surg 2017;402:1015–22. https://doi.org/10.1007/s00423-017-1593-6. Nov.
[55] Matics T, Khan N, Jani P, Kane J. The fragility index in a Cohort of Pediatric randomized controlled trials. J Clin Med 2017;14:6. https://doi.org/10.3390/jcm6080079. Aug, pii: E79.
[56] Brown J, Lane A, Cooper C, Vassar M. The results of randomized controlled trials in emergency medicine are frequently fragile. Ann Emerg Med 2018;Dec(18):S0196–644. https://doi.org/10.1016/j.annemergmed.2018.10.037. 31425-2.
[57] Tignanelli C, Napolitano L. The fragility index in randomized clinical trials as a means of optimizing patient care. JAMA Surg 2018:1–6. https://doi.org/10.1001/jamasurg.2018.4318. Nov.
[58] Akl E, Briel M, You J, Sun X, Johnston B, Busse J, et al. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. BMJ 2012;344:e2809. https://doi.org/10.1136/bmj.e2809. May.
[59] Klek S, Sierzega M, Turczynowski L, Szybinski P, Szczepanek K, Kulig J. Enteral and parenteral nutrition in the conservative treatment of pancreatic fistula: a randomized clinical trial. Gastroenterology 2011;141:157–63.
[60] Gianotti L, Besselink M, Sandini M, Hackert T, Conlon K, Gerritsen A, et al. Nutritional support and therapy in pancreatic surgery: a position paper of the International Study Group on Pancreatic Surgery (ISGPS). Surgery 2018;164:1035–48.
[61] Goudie AC, Sutton AJ, Jones DR, Donald A. Empirical assessment suggests that existing evidence could be used more fully in designing randomized controlled trials. J Clin Epidemiol 2010;63:983–91. https://doi.org/10.1016/j.jclinepi.2010.01.022.
[62] Sutton AJ, Cooper NJ, Jones DR. Evidence synthesis as the key to more coherent and efficient research 2009;9:1–9. https://doi.org/10.1186/1471-2288-9-29.
[63] Clarke M, Hopewell S, Chalmers I. Reports of clinical trials should begin and end with up-to-date systematic reviews of other relevant evidence: a status report. J R Soc Med 2007;100:187–90. Apr.
[64] Clarke M, Anderson P, Chalmers I. Discussion sections in reports of controlled trials published in general medical journals Islands in search of continents? J Am Med Assoc 2002;287:2799–801.
[65] Rosenthal R, Bucher H, Dwan K. The Use of systematic reviews when designing and reporting surgical trials. Ann Surg 2017;265:e35–6. https://doi.org/10.1097/SLA.0000000000001092. Apr.
[66] Turner R, Bird S, Higgins J. The impact of study size on metaanalyses: examination of underpowered studies in Cochrane reviews. PLoS One 2013;8.
[67] Pereira TV, Ioannidis JPA. Statistically significant meta-analyses of clinical trials have modest credibility and inflated effects. J Clin Epidemiol 2011;64:1060–9. https://doi.org/10.1016/j.jclinepi.2010.12.012.
[68] Wetterslev J, Jakobsen JC, Gluud C. Trial sequential analysis in systematic reviews with meta-analysis 2017:1–18. https://doi.org/10.1186/s12874-017-0315-7