Audit Clinical Evidence

Audit of Abdelbasset et al. (2019)

Three randomized controlled trials on exercise and depression

Published: 2026-09-13 **Modified:** 2026-09-13

Matthew B. Jané
University of Connecticut (Storrs, CT, USA)
matthew@auditclinicalevidence.org

Kutenda F. Mvududu
University of Massachusetts Boston (Boston, MA, USA)
kutenda@auditclinicalevidence.org

Report ID: `ACE2026001`

Audit of Abdelbasset et al. (2019)

Three randomized controlled trials on exercise and depression

Matthew B. Jané Kutenda F. Mvududu

2026-09-13

This report audits three randomized controlled trials by Abdelbasset and colleagues (2019a, 2019b, 2019c) evaluating exercise interventions for depression in patients with congestive heart failure. The three articles appear to report overlapping arms from the same underlying clinical trial while presenting them as separate studies. Across the articles, duplicated baseline characteristics and PHQ-9 outcome statistics indicate reuse of the same trial data even though the articles report incompatible authorship contribution statements and three distinct ethical approval numbers, and do not cite one another. The articles are potentially linked to a well-documented authorship-for-sale publication network. Furthermore, one of the three articles identifies Abdelbasset as both first author and editor which is a clear conflict of interest. Reported summary statistics are impossible given a real data set, specifically, seven of nine PHQ-9 (the primary outcome) mean and standard deviation pairs fail GRIMMER consistency tests. The reported outcomes also imply extreme between-group standardized mean differences of approximately three standard deviations at post-intervention, making the studies highly influential outliers in later meta-analyses. In total, this audit identifies 21 instances of Research Integrity Violations across seven categories. We conclude that the three articles are not credible and are unlikely to contain any evidentiary value thus we recommend retraction by the publishing journals.

Sections

1 Executive summary

1.A Description of audited studies

In this report, we investigate three studies: Abdelbasset (2019a), Abdelbasset (2019b), and Abdelbasset (2019c). Each of these studies report a randomized controlled trial (RCT) on the effectiveness of moderate-low intensity exercise on depression scores in patients with Congestive Heart Disease/Failure (CHF) related depression. In each study, moderate and low intensity exercise conditions consist of a 12-week intervention with three sessions per week of walking on a treadmill until subjects reached a specified maximum heart rate. Depression was measured using the Patient Health Questionnaire 9 (PHQ-9) at baseline, 6 weeks, and 12 weeks (i.e., post-intervention) for each study.

1.B Impact of audited studies

We describe a work that cites one of the audited studies as *infected*: its findings incorporate results that this report identifies as unreliable. Based on OpenAlex (retrieved 2026-09-08), the three studies are cited by 74 articles in total (infected articles). 17 of these are meta-analyses or reviews (infected reviews), and the infected reviews have themselves been cited 816 times, which is a measure of how far the audited results have spread through the evidence base; 22 clinical guidelines or consensus statements cite either an audited study or an infected review (see Table 1). The audited RCTs appear as the largest effects in the infected meta-analyses (e.g., Banyard 2025), with standardized mean differences over −2.9. A particularly high-impact infected meta-analysis with over 550 citations, Heissel (2023), reported an effect of −2.91 for Abdelbasset (2019a), making it the third largest effect in that analysis[1]. An effect of this magnitude has high leverage on the pooled estimate, so the accuracy and integrity of this effect is critical to evidence syntheses in this field.

Table 1. Citation counts for audited studies

Study	Infected articles	Infected reviews	Spread via reviews	Spread to guidelines
Abdelbasset (2019a)	37	6	80	3
Abdelbasset (2019b)	14	5	266	9
Abdelbasset (2019c)	23	6	470	10
Total	**74**	**17**	**816**	**22**

Note. Counts are from OpenAlex, retrieved 2026-09-08, and depend on its coverage; other databases (Scopus, Web of Science, Google Scholar) give different totals. A work that cites an audited paper is described here as infected, meaning its findings incorporate that paper's results. *Infected articles* is the number of works indexed by OpenAlex that cite the paper. *Infected reviews* is the subset of those works that OpenAlex types as a review or whose title indicates a meta-analysis or review. *Spread via reviews* is the total number of citations received by the infected reviews in that row, as a measure of how far the paper's results have propagated; a review that cites more than one audited paper is counted in each row, and a work citing several of those reviews is counted once per review, so rows and the total can overlap. *Spread to guidelines* is the number of distinct works whose title indicates a clinical guideline, consensus, or position statement and that cite the paper directly or cite one of its infected reviews; only guidelines published in journals are indexed, so policy documents from agencies such as WHO or NICE are not counted, and a guideline reached from more than one audited paper appears in each row.

[1] the two larger effects both have integrity issues which will be discussed in future reports.

2 | Authorship and publication integrity issues

Findings in this report are labeled using version 1.0 of the ACE Research Integrity Violations taxonomy (ACE-RIV). Section 1 of the taxonomy concerns the *Authorship and Publication* process. The following sections document the Section 1 RIVs found in the audited studies, along with the evidence for each one, and close with an ethical oversight issue (Section 3 of the taxonomy) that follows directly from the duplication described first.

2.A Duplicated data across studies

The three RCTs do not appear to represent three independent clinical trials. Rather, they appear to report overlapping and duplicated data from the same clinical trial.

The clearest evidence comes from the statistical tables displaying the differences between the exercise and control arms. Across the relevant articles, the baseline characteristics for these groups are identical, including age, sex distribution, BMI, ejection fraction, VO2peak, and all other clinical variables (see Figure 1). The reported PHQ-9 outcome values are also identical across the shared arms. The exercise group is reported with the same baseline, 6-week, and 12-week PHQ-9 scores, and the control group is likewise reported with the same baseline, 6-week, and 12-week values (see Figure 2).

The *Clinics* article reports a two-arm trial with 46 participants, consisting of a low-to-moderate exercise group and a control group. The later *Medicine* article reports a three-arm trial with 69 participants, reproducing the same low-to-moderate exercise and control-arm data while adding a third moderate-intensity exercise group. A separate *Medicine* article reports the same moderate-intensity exercise group as its own trial. None of the three articles cites the other two or mentions that the same arms appear elsewhere. The two smaller articles contain no participants or arms that are not in the three-arm article, and their baseline and PHQ-9 values are identical to it, so separate publication added nothing. This is consistent with duplicated publication or salami slicing, where the same underlying trial is split across multiple publications without transparent disclosure.

Accordingly, this report identifies two instances of **Duplicate Publication and Salami Slicing (RIV 1.5)**, corresponding to the two smaller articles (Abdelbasset 2019a, 2019c) whose findings are subsumed by the three-arm *Medicine* article (Abdelbasset 2019b).

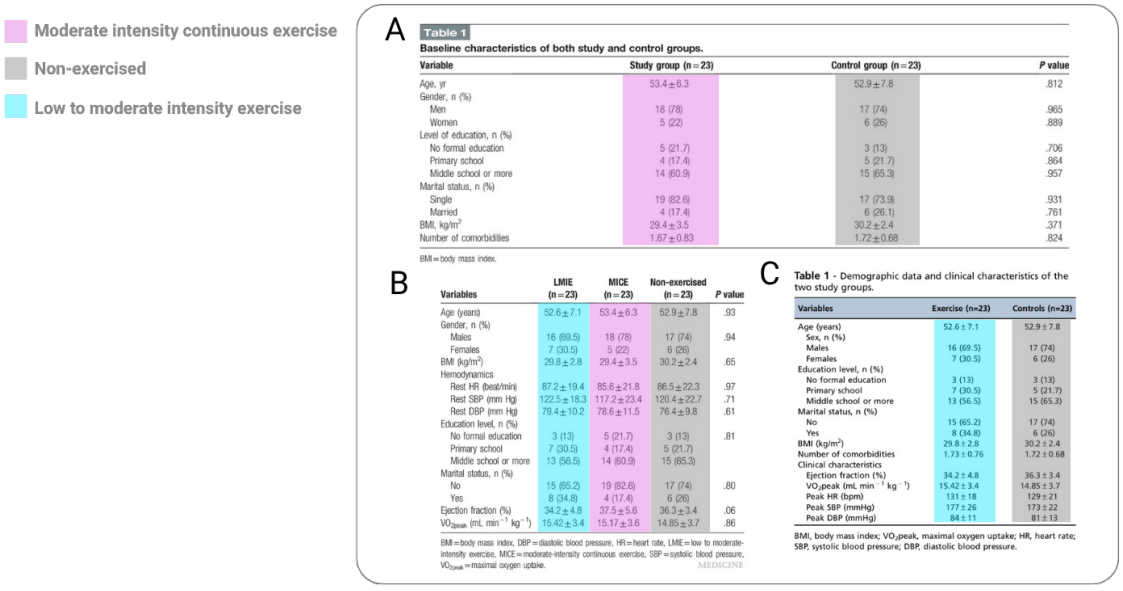

Figure 1. Baseline statistics for each of the three studies: panel A is Abdelbasset (2019a), panel B is Abdelbasset (2019b), panel C is Abdelbasset (2019c). Highlighted colors shows the matched data across studies.

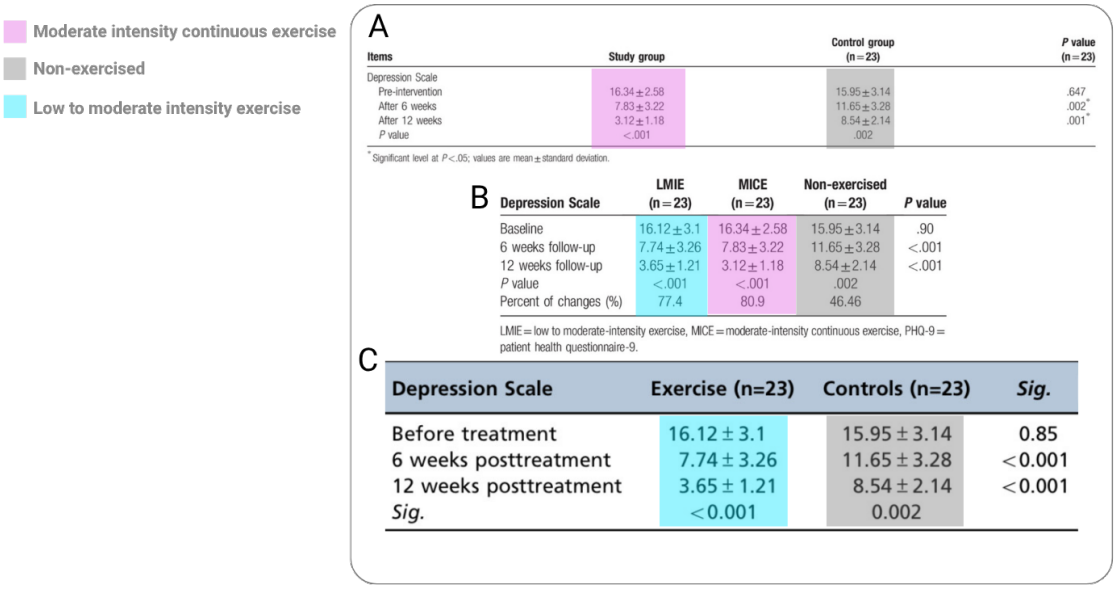

Figure 2. Statistics for PHQ-9 results (primary outcome) for each of the three studies: panel A is Abdelbasset (2019a), panel B is Abdelbasset (2019b), panel C is Abdelbasset (2019c). Highlighted colors shows the matched data across studies.

2.B Authorship-for-sale scheme

These three publications exhibit multiple characteristics consistent with a documented publication network that has previously been linked to authorship-for-sale schemes, which is when individuals pay for their names to be added as authors on scientific papers, often produced by "paper mills".

Independent investigations by Wise (2022) identified Walid Kamal Abdelbasset as a recurring member of a broader network of authors publishing across unrelated disciplines, including nanotechnology, chemistry, engineering, and aquaculture, despite Abdelbasset's primary appointment in physical therapy and rehabilitation. Multiple publications involving this network were directly matched to contemporaneous Facebook advertisements offering authorship positions for sale. For example, one publication involving Abdelbasset and members of this network titled, *Role of alloying composition on the nanomechanical behavior of amorphous nanolaminates*, was later retracted after the publisher concluded that there was sufficient evidence that several authorship positions had been purchased (Wise 2022).

Importantly, Wise (2022) emphasizes that the sale of authorship shows common indicators such as implausible collaborations, abrupt increases in publication volume, recurring author networks spanning unrelated disciplines, and, where available, advertisements matching published papers. The three RCTs examined here are consistent with these characteristics.

These publications were produced during the same period in which Abdelbasset's publication output started to increase dramatically. Wise identified Abdelbasset as one of several authors exhibiting an abrupt rise in publication volume consisting of papers that were often substantively unrelated to his discipline. Wise described this pattern as the emergence of modern "Renaissance men and women," in which the same authors repeatedly appeared on papers spanning numerous unrelated scientific disciplines (Wise 2022). An analysis of Scopus showed that Abdelbasset published in fields such as mechanical properties of amorphous metallic glasses, sulfone synthesis and applications in organic chemistry, probiotic applications in aquaculture health, inequalities and means in mathematical functions, and modified atmosphere packaging for fresh produce quality. This is highly unusual for an academic specializing in physiotherapy and women's health.

Taken together, these indicators, along with the additional issues documented in what follows as well as the data duplication and salami slicing described in Section 2.A, warrant three instances of **Authorship for Sale and Paper Mill Activity (RIV 1.3)**, one per article. This classification does not depend on evidence that

authorship on these three articles was sold; we have none. It depends on the documented link between the authors and an authorship-for-sale network, combined with problems in the articles themselves (duplicated trial data, incompatible contribution statements, three ethics approvals, self-editing, and impossible summary statistics) that are characteristic of that kind of publishing.

2.C Compromised peer review process due to self-editing

The published article for Abdelbasset (2019a) identifies Walid Kamal Abdelbasset as the editor of the same study on which he is also the first author. This represents a serious violation of publication ethics. Whoever reviewed the manuscript, the person who chose the reviewers and decided on their reports was the first author, so the review was not independent. Because an author cannot independently oversee peer review of his own manuscript, this report classifies the issue as one instance of **Peer Review Subversion (RIV 1.7)**.

This self-editing issue adds additional support to the three instances of **Authorship for Sale and Paper Mill Activity (RIV 1.3)** described previously in Section 2.B. Compromised editorial handling is a known method by which paper mill and authorship-for-sale publications can enter the literature.

It is worth noting that two of the three studies (Abdelbasset 2019a, 2019c) were both published in the journal *Medicine* where Abdelbasset is on the editorial board for physiotherapy and rehabilitation (Medicine n.d.).

2.D Incompatible authorship contribution statements

The authorship contribution statements are incompatible with the conclusion that these papers report independent studies. The two heart-failure papers appear to report the same underlying trial data, but they assign materially different contributors to core research roles that should not change if the underlying study is the same.

In the two-arm *Medicine* article (Abdelbasset 2019a), conceptualization is attributed only to Abdelbasset; data curation and formal analysis are attributed only to Abdelbasset and Alqahtani; funding acquisition is attributed only to Abdelbasset; and resources and software are attributed only to Alqahtani. By contrast, the later three-arm *Medicine* article assigns conceptualization to Abdelbasset, Alqahtani, and Ahmed; data curation to five authors; formal analysis to Abdelbasset, Ahmed, and Ibrahim; funding acquisition to Alqahtani, Alrawaili, Ahmed, and Ibrahim; resources only to Ibrahim; and software to Alrawaili and Elnegamy. The same reported trial cannot coherently have one set of individuals who acquired funding, curated the data, performed the analysis, supplied resources, and provided software in one article, and a different set of individuals performing those same roles in another article.

The *Clinics* article's (Abdelbasset 2019c) author contribution statement assigns study design, data curation, and analysis primarily to Abdelbasset, Alqahtani, Elshehawy, Tantawy, Elnegamy, and Kamel. Several authors credited with core roles in the three-arm *Medicine* article are absent from the *Clinics* contribution statement, while several authors credited in the *Clinics* article are absent from the three-arm *Medicine* article. This is difficult to reconcile if the papers derive from the same trial data.

Because no article acknowledges the other two, none explains why different people are credited with the same core roles. These incompatible contribution statements warrant three instances of **Incompatible Authorship or Contribution Statements (RIV 1.10)**, one per article, because the three studies indicate that authorship roles were not assigned according to stable, verifiable contributions to the underlying clinical trial.

2.E Incompatible ethical approval numbers

The ethical approval statements also differ across the three articles. The two-arm *Medicine* article reports ethical approval number **RHPT/017/008** (Abdelbasset 2019a). The *Clinics* article reports ethical approval number **RHPT/017/009** (Abdelbasset 2019c). The three-arm *Medicine* article reports ethical approval number **RHPT/017/010** (Abdelbasset 2019b).

This discrepancy is difficult to reconcile with the duplicated trial data described in Section 2.A. If the three articles reported three independent clinical trials, different ethical approval numbers would not necessarily be problematic. However, the articles do not appear to report independent trials. The *Clinics* article reproduces the same low-to-moderate intensity exercise and control-arm data that appear in the later three-arm *Medicine* article, while the separate two-arm *Medicine* article reports the moderate-intensity exercise comparison as its own trial. Thus, the same apparent underlying clinical trial is associated with three different ethical approval numbers.

The structure of the approval numbers suggests a common institutional or committee prefix (**RHPT**, probably referring to rehabilitation and physical therapy), a shared year or protocol field (**017**, probably corresponding to 2017), and a final sequential approval identifier (**008**, **009**, **010**). The three articles therefore do not report differently formatted versions of the same approval number. They report three consecutive but distinct approval identifiers.

This inconsistency is especially concerning because the reported trial is within a highly vulnerable population. The participants had chronic heart failure, were assigned to exercise interventions, and had high baseline PHQ-9 scores indicating substantial depressive symptom burden. In such a context, clear ethical approval is particularly important. Without clarification, the presence of three different approval numbers for overlapping trial data raises serious concerns about the adequacy of the ethical oversight.

No article mentions the approval numbers reported in the other two, and none explains how one trial could carry three approvals. Accordingly, without further clarification by the authors this report classifies the issue as three instances of **Ethics Approval Anomalies (RIV 3.3)**, one per article, because the ethical approval information is not reconcilable with the apparent reuse of the same underlying trial data across multiple publications.

3 Impossible data and extreme treatment effects

3.A Impossible means and standard deviations

The results table in each of the three RCTs show means and SDs of PHQ-9 scores across each group for the baseline, 6-week, and 12-week timepoints. Since PHQ-9 is scored as a sum of item responses and therefore the resulting total score is on an integer scale. Using the GRIM/MER test (Brown 2017; Anaya 2016), we can demonstrate that the means and standard deviations are impossible to be produced given the sample size. In all three RCTs, it is made clear that each group has 23 patients and none of them were lost to follow-ups at 6 or 12 weeks. Table 2 shows the extracted PHQ-9 mean, standard deviation, and sample size values for each of the three groups (LMIE, MICE, Non-Exercised) and the three time-points (baseline, 6-weeks, 12-weeks), making $3 \times 3 = 9$ testable GRIMMER cases. The results of the GRIMMER tests can be seen in Table 2. Since Abdelbasset (2019b) reports all groups in their results, we extract the summary statistics from the corresponding table (panel B of Figure 2).

The results displayed in Table 2 show that of the nine GRIMMER cases, seven of them are inconsistent. Five of the seven cases that are inconsistent have means that can not possibly be calculated from a real data set (i.e., GRIMMER inconsistent) and the remaining two cases show standard deviations that can not come from a possible data set (i.e., GRIMMER inconsistent).

Table 2. GRIMMER consistency for reported PHQ-9 means and standard deviations.

Timepoint	Group	Mean	SD	n	Consistent	Reason
Baseline	LMIE	16.12	3.10	23	No	GRIM inconsistent
Baseline	MICE	16.34	2.58	23	No	GRIM inconsistent
Baseline	Control	15.95	3.14	23	No	GRIM inconsistent
6 weeks	LMIE	7.74	3.26	23	Yes	Passed all
6 weeks	MICE	7.83	3.22	23	No	GRIMMER inconsistent (test 3)
6 weeks	Control	11.65	3.28	23	Yes	Passed all
12 weeks	LMIE	3.65	1.21	23	No	GRIMMER inconsistent (test 3)
12 weeks	MICE	3.12	1.18	23	No	GRIM inconsistent
12 weeks	Control	8.54	2.14	23	No	GRIM inconsistent

textitNote. SD = standard deviation, n = sample size, LMIE = Low to Moderate Intensity Exercise, MICE = Moderate-Intensity Continuous Exercise, Control = Non-exercised

One possible explanation for the GRIMMER failures is that the reported group sample size of $n = 23$ was incorrect. This explanation is unlikely. The group sample sizes are reported consistently within the articles, and the same $n = 23$ group sizes appears across the study articles. Nevertheless, we evaluated whether the inconsistencies could be resolved by assuming smaller analyzed sample sizes, such as might occur if some participants were excluded from the final analysis.

To test this, we recalculated GRIMMER consistency for every candidate group sample size from $n = 2$ through the reported $n = 23$. For each group and each candidate sample size, Figure 3 shows whether each of the three reported PHQ-9 mean-SD pairs (one per timepoint) is GRIMMER-consistent. A plausible sample-size correction would require all three timepoints within a group to become consistent at the same candidate sample size, which would appear in the figure as a fully filled column. No such column exists for any group at any sample size from $n = 2$ to $n = 23$.

Thus, the GRIMMER failures are not resolved by assuming that the true sample sizes were smaller than reported. The inconsistencies instead persist across the possible smaller group sizes.

Attrition does not resolve them either. If participants had dropped out over the course of the trial, the analyzed sample sizes within a group would decline across timepoints, so that the baseline sample size is at least the 6-week sample size, which is at least the 12-week sample size. In Figure 3, a possible attrition pattern would appear within a group as one filled square in each row, at sample sizes that do not increase from Baseline down to 12 weeks. No such pattern exists in any group. In particular, the 12-week means and SDs are GRIMMER-inconsistent at every candidate sample size from $n = 2$ to $n = 23$ in all three groups, so there is no number of analyzed participants at 12 weeks, with or without attrition, that could have produced the reported values; the MICE baseline pair is likewise inconsistent at every candidate sample size.

These failures cannot be explained away by rounding, by the scale, or by missing data. GRIMMER already allows for means and standard deviations rounded to two decimals, the PHQ-9 is an integer score between 0 and 27, the articles report no loss to follow-up, and no smaller sample size reproduces the values. This is classified as seven instances of **Statistical Inconsistencies and Impossible Data Sets (RIV 2.2)**, one for each failed GRIMMER test.

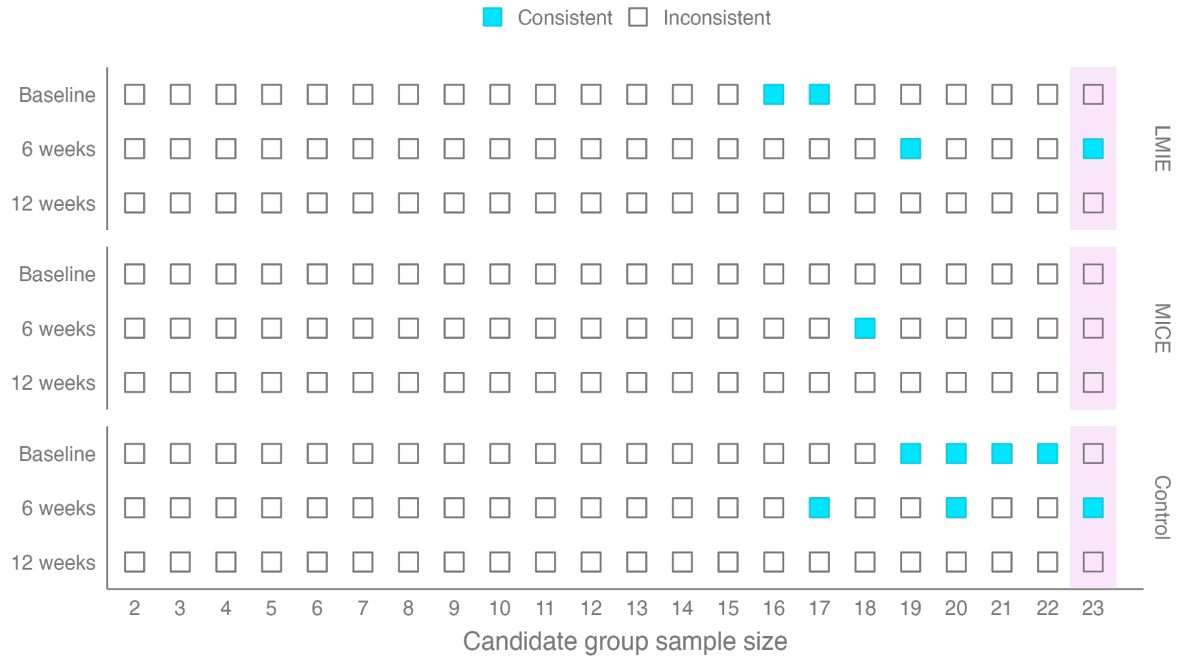

Figure 3. GRIMMER consistency of the reported PHQ-9 means and SDs for every candidate group sample size from $n = 2$ to the reported $n = 23$. Each row is one timepoint within a group; a filled cyan square marks a mean-SD pair that is GRIMMER-consistent at that sample size, a hollow square one that is not. The shaded column is the reported sample size. A sample-size error would show as a column with all three squares filled within a group; no such column exists.

3.B Extreme treatment effects

The reported PHQ-9 outcomes imply extremely large treatment effects. In meta-analyses that include at least one of the three Abdelbasset studies, the Abdelbasset effect is consistently among the largest effects in the meta-analysis. For example, see the meta-analyses by Bricca (2020) and Banyard (2025). Across these meta-analyses, the reported effects are generally near −2.90, with some variation depending on the estimation procedure and the treatment arm used for the comparison.

Using the values reported in the three-arm article, this report estimates the 12-week between-group standardized mean differences as −2.81 for the low-to-moderate intensity exercise group and −3.14 for the moderate-intensity continuous exercise group, each compared with the non-exercised control group. These values imply that, after 12 weeks, the exercise groups had PHQ-9 scores approximately three standard deviations lower than the control group.

These effects are implausibly large in the context of clinical trials. According to the benchmarks proposed by Sawilowsky (2009), standardized mean differences above 0.80 are considered large, values above 1.20 are considered very large, and values above 2.00 are considered huge. The reported Abdelbasset effects exceed even the “huge” benchmark by approximately one entire standard deviation. Although effect-size benchmarks are only descriptive heuristics and cannot by themselves establish an error, effects of this magnitude warrant explanation, particularly when combined with the duplicated publication pattern (Section 2.A), GRIMMER failures (Section 3.A), and other internal inconsistencies documented in this report.

The articles do not comment on the size of these effects, compare them with earlier trials, or point to anything about the intervention or the sample that would account for them. A reporting error does not explain them either: the standardized mean differences are recomputed from the means, standard deviations, and group sizes printed in the three-arm article, and the two smaller articles print the same values. This warrants two instances of **Anomalies and Extreme Values (RIV 2.1)** for the two extreme between-group effects.

3.C Surprising level of treatment compliance

The three Abdelbasset articles report that all randomized participants were retained through the 12-week follow-up. In all three arms of the underlying trial, 69 patients were randomized into three groups of 23, and all three groups were analyzed at follow-up (Abdelbasset 2019b). Thus, the demands of the intervention, the reported trial record implies zero attrition.

This is not impossible, but it is surprising. The participants were patients with congestive heart failure and moderate-severe depressive symptoms. The intervention was not passive. Participants assigned to exercise were expected to complete treadmill exercise three times per week for 12 consecutive weeks, with intensity adjusted according to maximum heart rate. A study population with heart failure and depression would ordinarily be expected to include at least some missed complications, withdrawals, or incomplete follow-up assessments.

The concern is strengthened by the limited reporting of adherence information. The articles state that the exercise groups were supervised and received feedback to adjust intensity, but they do not provide session attendance, missed-session counts, adherence percentages, reasons for missed visits, adverse-event monitoring by group, or the number of participants who completed the full exercise regimen.

Perfect retention is another irregularity that warrants scrutiny. The absence of attrition is difficult to evaluate independently because the articles do not provide the participant-level adherence or follow-up information needed to verify the reported trial conduct.

4 Conclusion

The three Abdelbasset et al. articles audited in this report contain multiple anomalies and errors across publication, authorship, ethics approval, and statistical reporting. The articles appear to report overlapping trial arms from the same underlying clinical trial across multiple publications, without transparent disclosure that the articles are linked in any way. This pattern supports two instances of **Duplicate Publication and Salami Slicing (RIV 1.5)**.

The publication record also contains serious authorship and editorial concerns. The studies are connected to an author network previously associated with authorship-for-sale concerns, and one article identifies Abdelbasset as both first author and editor of the same manuscript. The contribution statements across the articles are also incompatible, assigning different individuals to core study roles that should be the same if the same clinical trial is being reported. These issues support three instances of **Authorship for Sale and Paper Mill Activity (RIV 1.3)**, one instance of **Peer Review Subversion (RIV 1.7)**, and three instances of **Incompatible Authorship or Contribution Statements (RIV 1.10)**.

The ethical approval is also not reconcilable with the duplication of trial data. The three articles report three distinct ethical approval numbers, despite reporting overlapping arms from the same clinical trial. This is especially concerning because the reported trial population was highly vulnerable: participants had chronic heart failure, were assigned to exercise interventions, and had high baseline depressive symptoms. In this context, clear ethical oversight and approval is essential. Without further clarification, the differing approval numbers warrant three instances of **Ethics Approval Anomalies (RIV 3.3)**.

The statistical evidence raises severe concerns about the integrity of the data. The summary statistics for the PHQ-9 outcome consistently fail GRIMMER checks, indicating that the reported summary statistics could not have been produced by a real data set (this warrants seven instances of **Statistical Inconsistencies and Impossible Data Sets (RIV 2.2)**). These failures are not resolved by when allowing for possible unreported list-wise deletion of some individuals from the data set. The reported outcomes also imply extremely large between-group treatment effects, approximately three standard deviations in magnitude at 12 weeks, which are far beyond conventional benchmarks for large treatment effects and appear as

highly influential outliers in later meta-analyses. Due to the massive effect size observed in the articles it warrants two instances of **Anomalies and Extreme Values (RIV 2.1)**.

Taken together, these findings substantially undermine the reliability and integrity of the three reported RCTs. This report cannot verify that the reported trial data correspond to an actual conducted randomized controlled trial as described in the publications. The combination of duplicated trial arms, incompatible authorship roles, inconsistent ethical approval identifiers, impossible descriptive statistics, and extreme treatment effects leaves the provenance of the reported patient data unresolved.

We believe that the appropriate editorial action is *retraction*. The journals that published these articles, *Medicine* and *Clinics*, should retract all three articles because they contain duplicated trial arms, incompatible ethical approval identifiers, incompatible authorship contribution statements, compromised editorial handling, impossible PHQ-9 results, extreme treatment effects that have influenced later meta-analyses, and are potentially a product of a well-documented authorship-for-sale publication mill. These are mutually reinforcing failures that undermine the integrity and evidentiary value of the reported clinical trial.

Until retractions are issued, meta-analysts, guideline developers, and readers should treat the three studies as unreliable and exclude them from evidence syntheses.

Table 3. Summary of Research Integrity Violations identified across the audited Abdelbasset et al. studies (ACE-RIV v1.0)

RIV	Title	Instances	Section
RIV 1.3	Authorship for Sale and Paper Mill Activity	3	Section 2.B
RIV 1.5	Duplicate Publication and Salami Slicing	2	Section 2.A
RIV 1.7	Peer Review Subversion	1	Section 2.C
RIV 1.10	Incompatible Authorship or Contribution Statements	3	Section 2.D
RIV 2.1	Anomalies and Extreme Values	2	Section 3.B
RIV 2.2	Statistical Inconsistencies and Impossible Data Sets	7	Section 3.A
RIV 3.3	Ethics Approval Anomalies	3	Section 2.E
Total		**21**	

Version history

Table 4. Any change to a finding, RIV count, or recommendation creates a new version with a new version DOI. Superseded versions remain permanently resolvable.

Version	Date	DOI	Change
1	2026-09-13	reserved at deposit	Initial publication.

Abdelbasset, Alqahtani (2019a) A randomized controlled trial on the impact of moderate-intensity continuous aerobic exercise on the depression status of middle-aged patients with congestive heart failure. *Medicine*. DOI: 10.1097/MD.0000000000015344.

Abdelbasset, Alqahtani, Alrawaili, Ahmed, Elnegamy, Ibrahim, Soliman (2019b) Similar effects of low to moderate-intensity exercise program vs moderate-intensity continuous exercise program on depressive disorder in heart failure patients: A 12-week randomized controlled trial. *Medicine*. DOI: 10.1097/MD.0000000000016820.

Abdelbasset, Alqahtani, Elshehawy, Tantawy, Elnegamy, Kamel (2019c) Examining the impacts of 12 weeks of low to moderate-intensity aerobic exercise on depression status in patients with systolic congestive heart failure-a randomized controlled study. *Clinics*. DOI: 10.6061/clinics/2019/e1017.

Banyard, Edward, Garvey, Stephenson, Azevedo, Benson (2025) The effects of aerobic and resistance exercise on depression and anxiety: Systematic review with meta-analysis. *International Journal of Mental Health Nursing*. DOI: 10.1111/inm.70054.

Heissel, Heinen, Brokmeier, Skarabis, Kangas, Vancampfort, Stubbs, Firth, Ward, Rosenbaum (2023) Exercise as medicine for depressive symptoms? A systematic review and meta-analysis with meta-regression. *British Journal of Sports Medicine*. DOI: 10.1136/bjsports-2022-106282.

Wise, Magazinov (2022) The incredible collaborations of renaissance men and women. *For Better Science*. URL: https://forbetterscience.com/2022/10/19/the-incredible-collaborations-of-renaissance-men-and-women/.

Wise (2022) Comments on ”role of alloying composition on the nanomechanical behavior of amorphous nanolaminates”. URL: https://pubpeer.com/publications/01530E72364AB34E7CFC9F3A1ADD86#1.

Medicine (n.d.) Editorial board. URL: https://journals.lww.com/md-journal/Pages/editorialboard.aspx#:~:text=Walid%20Kamal%20Abdelbasset.

Brown, Heathers (2017) The GRIM test: A simple technique detects numerous anomalies in the reporting of results in psychology. *Social Psychological and Personality Science*. DOI: 10.1177/1948550616673876.

Anaya (2016) The GRIMMER test: A method for testing the validity of reported measures of variability. *PeerJ*. DOI: 10.7287/peerj.preprints.2400v1.

Bricca, Harris, Jäger, Smith, Juhl, Skou (2020) Benefits and harms of exercise therapy in people with multimorbidity: A systematic review and meta-analysis of randomised controlled trials. *Ageing Research Reviews*. DOI: 10.1016/j.arr.2020.101166.

Sawilowsky (2009) New effect size rules of thumb. *Journal of Modern Applied Statistical Methods*. DOI: 10.22237/jmasm/1257035100.