Anemia in Placebo Arms of Cancer Studies

JOHANNES E. WOLFF

AbbVie Inc, North Chicago, IL, U.S.A.

Abstract. *Background/Aim: Adverse event (AE) frequencies observed in interventional clinical trials are difficult to interpret when the placebo control is missing. Materials and Methods: Systematic literature review of AEs reported from the placebo arms of randomized cancer trials between 2008 and 2021. Imputation of missing values assuming normal distribution of hemoglobin values. Results: Anemia grade 1 or higher was reported in 46 of 100 placebo monotherapy cohorts with a mean frequency of 23.4% (SD=27%) of the enrolled patients. The reported frequency depended on the type of cancer; other demographic variables had no significant influence on anemia frequency. Conclusion: External controls for anemia in clinical trials should be disease specific.*

Clinical trials without control group are common in early oncology drug development. The frequency of adverse events (AEs) observed in these studies may be compared to bench marks, which may be provided by placebo arms of randomized trials (1, 2). However, outcome data also depend on the demographics of the patient cohort, and bench marks might need to be adjusted. Using the example of headache, a previous analysis showed a counterintuitive relation of more headache reported in healthier patients (described as ECOG status) (3). This project addresses anemia, as an example of an AE that is based upon more objective laboratory values.

Adverse events are defined independent of causality, but in clinical trials they are typically described as treatment emergent AEs (TEAEs), defined as occurring during treatment – not at diagnosis. Temporality is one of the key variables to determine causality (4, 5) and therefore TEAEs are an interim between all AEs and adverse drug reactions (ADRs), and AE frequencies observed as symptoms of the diagnosis may not be appropriate external controls for

TEAEs. For instance, thrombocytopenia is a typical laboratory finding of leukemia. However, in clinical trials only newly occurring or worsening thrombocytopenia during treatment will be listed as TEAE, resulting in potentially lower reported frequencies in AML than other cancer diagnoses. This project addresses TEAE in the context of clinical trials.

AEs are typically named using preferred terms (PTs) listed in medical dictionary for regulatory activities (MEDRA), and graded as defined by the National Cancer Institute as Common Terminology Criteria for Adverse Events (CTCAE). CTCAE version 5 combines laboratory values and clinical information for anemia grading: grade 1: Hemoglobin (Hgb) <lower level of normal (LLN) - 10.0 g/dl, grade 2: Hgb <10.0-8.0 g/dl, grade 3: Hgb <8.0 g/dl transfusion indicated, grade 4: Life-threatening consequences; urgent intervention indicated, and grade 5: Death. The concept and the numeric limits have changed over the various versions since its creation in 1983 (6). Grade 5 was not listed in version 2. In versions 2 and 3 the word "Hemoglobin" was used instead of "Anemia"; and the limit between grade 3 and grade 4 was defined by a numeric laboratory value (Hgb <6.5 g/dl); Version 4 was identical to version 5. The complexity of CTCAE reporting has created a burden for cooperative clinical trials with limited funding (7); and the adherence to the CTCAE definitions was not always optimal. In particular, in studies without control groups, adherence to CTCAE reporting was low (8).

AEs observed in placebo arms of randomized trials are potential sources for external control data allowing a better assessment of outcomes of interventional trials. Meta-analyses and systematic literature reviews of AE literature reviews have recently contributed to drug development, and the methodology is developing (1, 2). Interestingly, there was a high level of heterogeneity between placebo arms of different studies, contrasted by a close correlation between in placebo arms and treatment arms within the same study (2). Both of these findings could reflect the influence of the eligibility criteria and patient demographics (9). A previous similar analyses of AEs in placebo arms (2, 3) showed the frequency of headache reporting to be higher among studies with high proportion of ECOG 0 performance status. Here, the hypothesis is addressed that anemia reporting was related

Correspondence to: Dr. med. habil. Johannes E. Wolff, MD, Ph.D., 1709 Napa Suwe Ln, Wauconda, IL 60084, U.S.A. Tel: +1 2242216093, Fax: +1 8479380660, e-mail: johannesewolff@gmail.com
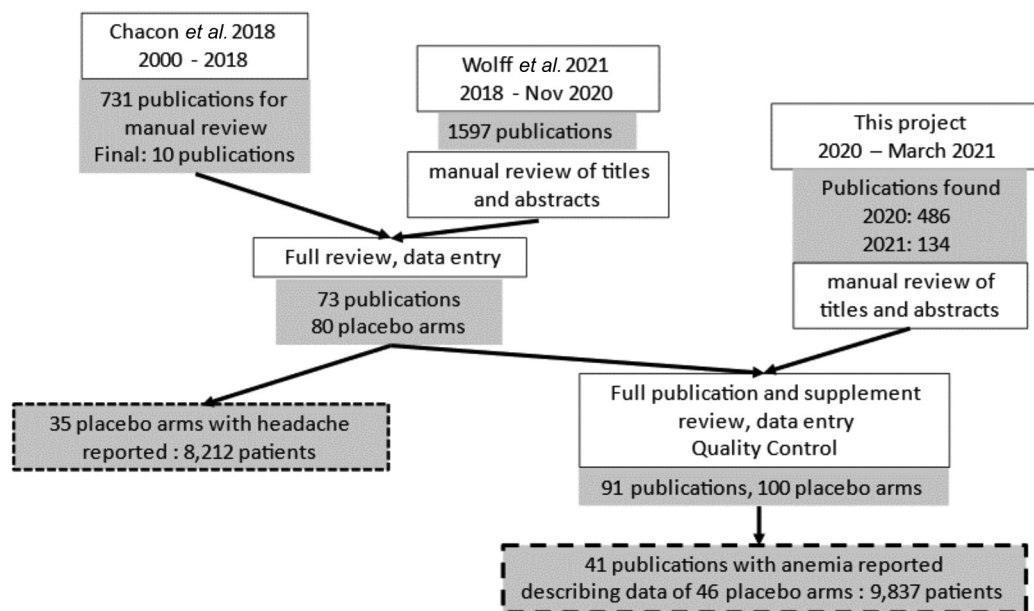
Figure 1. *The data collection includes searches done for previous publications (2, 3). For the anemia analysis presented here, the search was updated to include the complete year 2020 and the beginning of 2021. The specific search and the final list of included publications is available upon request.*

to different demographic variables than headache; and bench marks are described. The data collection was expanded, and model based imputation for missing values was added.

Materials and Methods

The selection of included publications built upon two previous meta-analyses 2000-2018 (2), and 2018-Nov 2020 (3). In addition, the search was repeated and expanded for 2020 until March 2021 in PubMed using the same method. The search combined 1) cancer, 2) randomized and 3) AE, with each of these terms described by multiple synonyms. The complete algorithm is available upon request. The new search identified 486 titles in 2020 and 134 in 2021, which eventually resulted in the inclusion of 91 publications (Figure 1). The search algorithm and list of included articles is available from the corresponding author upon request.

Anemia was reported in various ways with a total of 11 different definitions. For instance, commonly grade 1 or higher was reported or grade 3 and 4 together as one numeric value. Two imputation steps were used to fill in missing values. First, all values that could be logically imputed were calculated. For instance, if all of the grades 1-5 were provided individually, then grade 0 and all the combined grades (grade 1 or higher) could be calculated. The second imputation was model-based and was only done when at least one original data value other than grade 5 was reported. The model assumed normal distribution of the underlying Hgb values, started with specific values for mean and SD of the Hgb distributions for each cohort, which were then translated in frequencies of grades using the CTCAE definitions. The quality of a specific model was measured by the sum of differences between the known and the modelled values for each data line and anemia grade. The assumed values for mean and SD were then modified

using a random factor. The new model substituted the previous one, if it was found to match the raw data more accurately, and then the process was repeated. For each line of data, 1,500 repetitions of this loop were performed. After optimizing the model, original data were used to overwrite the calculated data when available; and finally, the sum of the % values for each grade was adjusted to 100% by proportionally adjusting the modelled values only. The programming code of the imputation will be provided by the corresponding author upon email request.

For categorical variables such as cancer indication, ANOVA was used to compare the frequency of AEs among subgroups. For quantitative variables, linear regression, Pearson's correlation and visual evaluation in scatter plots were used; and quantitative variables were transferred to categorical variables allowing the use ANOVA for multiple variables. This was first done after imputation for data from the placebo arms only, and then repeated as sensitivity analysis for the raw data prior to imputation, and then for all data, including the treatment arms of the studies. All analyses and *p*-values were exploratory. SPSS (Statistical Package for Social Studies, IBM version 23.0) was used to for the model and to conduct the analyses.

Results

The 91 included publications described 208 patient cohorts with a total of 47,962 patients. Among those were 100 placebo monotherapy arms describing 20,581 patients. The core of the analysis is built upon a subgroup: the placebo monotherapy cohorts that provided anemia data. Those included 41 publications reporting 46 patient cohorts with 9,837 patients. The majority of data were derived from phase

Table I. *Demographics*.

	All cohorts[1] (208 cohorts)	Placebo monotherapy[2] (100 cohorts)	Placebo+Tx with anemia data[3] (91 cohorts)	Placebo monotherapy with anemia data[4] (46 cohorts)
Number of patients, mean (range)	231 (5-2,256) (n=208)*	205.8 (10-2,253) (n=100)*	247.3 (10-1,944) (n=91)*	213.9 (10-1,944) (n=46)*
Year of Publication, mean (range)	2019 (2009-2021) (n=208)*	2019 (2009-2021) (n=100)*	2019 (2015-2021) (n=91)*	2019 (2015-2021) (n=46)*
Median age, mean (range)	56.4 (7.1-74.5) (n=199)*	56.8 (7.1-74) (n=96)*	58.6 (31-74) (n=87)*	59.2 (31-72) (n=44)*
Male sex (%)	52 (0-100) (n=192)*	52.5 (0-100) (n=92)*	41.5 (0-100) (n=89)*	42.6 (0-100) (n=45)*
ECOG 0, % (range)	58.1 (2.1-100) (n=154)*	57 (5.2-100) (n=76)*	54.8 (11.1-100) (n=73)*	54.9 (11.4-100) (n=38)*
Phase 1/2/3, % (range)	5.8/27.9/66.3 (n=208)*	4/27/69 (n=100)*	0/17.6/82.4 (n=91)*	0/17.4/82.6 (n=46)*
Cancer organ, % (n)				
Liver	10.6 % (22)	11% (11)	8.8% (8)	8.7% (4)
Hem	9.6% (20)	10% (10)	11% (10)	10.9% (5)
Breast	8.7% (18)	8% (8)	8.8% (8)	8.7% (4)
Lung	8.7% (18)	9% (9)	4.4% (4)	4.3% (2)
Colorectal	8.2% (17)	10% (10)	9.9% (9)	13% (6)
Ovary	7.7% (16)	8% (8)	17.6% (16)	17.4% (8)
Skin	5.8% (12)	6% (6)	2.2% (2)	2.2% (1)
No organ	5.8% (12)	4% (4)	0% (0)	0% (0)
Kidney	5.3% (11)	5% (5)	0% (0)	0% (0)
Prostate	4.8% (10)	5% (5)	4.4% (4)	4.3% (2)
Soft tissue	4.8% (10)	5% (5)	6.6 % (6)	6.5% (3)
Gastric	3.8% (8)	3% (3)	8.8% (8)	8.7% (4)
Uterus	2.9% (6)	2% (2)	6.6% (6)	4.3% (2)
Pancreas	1.9% (4)	2% (2)	4.4% (4)	4.3% (2)
Tumor status values, % (n)				
No Cancer[5]	34.1% (71)	32% (32)	19.8% (18)	19.6% (9)
Measurable[6]	29.8% (62)	30% (30)	28.6% (26)	26.1% (12)
Mixed status[7]	34.1% (71)	37% (37)	51.6% (47)	54.3% (25)
Line of treatment values, % (n)				
Newly diagnosed	39.9% (83)	38% (38)	33% (30)	30.4% (14)
R/R	41.8% (87)	45% (45)	53.8% (49)	56.5% (26)
Mixed new-R/R[8]	9.6% (20)	10% (10)	8.8% (8)	8.7% (4)
Healthy V.[9]	6.7% (14)	5% (5)	0% (0)	0% (0)
Prevention[10]	1.9% (4)	2% (2)	4.4% (4)	4.3% (2)
Line of previous Tx**	0.8 (0-4) (n=143)*	0.8 (0-4) (n=67)*	1.1 (0-4) (n=64)*	1.2 (0-4) (n=31)*
Route values, % (n)				
PO	75 (156)	76% (76)	75.8% (69)	76.1% (35)
IV	16.3% (34)	15% (15)	19.8% (18)	19.6% (9)
SC	4.8% (10)	5% (5)	2.2% (2)	2.2% (1)
IM	1.9% (4)	2% (2)	0% (0)	0% (0)
ID	1.9% (4)	2% (2)	2.2% (2)	2.2% (1)

*n, number of available values; **Data presented as median (range). [1]All lines of data regardless of treatment or anemia data available; [2]All cohorts treated with placebo only regardless if anemia data were reported; [3]All cohorts that had anemia data reported regardless of the treatment; [4]Patient cohorts treated only with placebo and with anemia data reported (core data of this analysis); [5]Studies enrolling only patients with no measurable disease; [6]Studies enrolling only patients with measurable disease; [7]Studies enrolling patients regardless of measurable disease; [8]Studies enrolling both newly diagnosed and relapsed or refractory disease; [9]Healthy Volunteers; [10]Cancer prevention studies enrolling patients at risk but without cancer diagnosis. R/R: Relapsed/refractory; IV: intravenous; PO: per os; SC: subcutaneous; IM: intramuscular; ID: intradermal; Tx: treatment.

3 studies between 2018 and 2021, the mean of median ages was 56.4 years, and the mean percentage of males was 52%. Further details of the demographics are described in Table I.

The grade of anemia was reported in different ways: The most common recorded value among the 100 placebo monotherapy cohorts was grade 5 anemia (48 cohorts), which was always 0. However, this value was commonly not listed in the adverse event table, but asserted from the text of the Results, where it was commonly mentioned that no death occurred as a result of an adverse event, or specifically listed the adverse events leading to death. The second most commonly reported value was grade 1 or higher (any anemia, reported in 21 cohorts), followed by grade 3 (17 cohorts), grade 4 (17 cohorts), grade 3 or higher (14 cohorts), grade 1or 2 (11 cohorts), grade 3 or 4 (11 cohorts), grade 0 (3 cohorts), grade 1 (3 cohorts), grade 2 (3 cohorts), grade 2 or higher (2 cohorts). Logical imputation allowed the calculation of further values increasing the number of numeric values for grade 3 or higher to 36 and grade 1 or higher to 35. Model-based imputation finally allowed 46 values for all classifications of grades. Among the 46 placebo cohorts (placebo monotherapy with anemia data), the mean of the calculated median Hgb values was 14.9 g/dl and the mean of the standard deviations was 2.4 mg/dl. All analyses that resulted in p-values lower than 0.05 among the raw data, also showed similar or lower p-values in the imputed data; and all analyses resulting in a p-value <0.05 after model supported imputation had at least a trend in the same direction in the raw data. However, more analyses reached such low exploratory p-values when calculated from the larger data set of imputed values, which were also the source for the bench marks reported below.

The mean frequency of anemia grade ≥1 reported in 46 of 100 placebo monotherapy cohorts was 23.4% (SD=27.0, range=0-100), grade ≥2 10.2% (SD=19.7, range=0-76.3), grade ≥3 5.6% (SD=9.7, range=0-46); and for grade 5 it was 0%. When normalizing these numbers to the numbers of patients included in cohort, studies with large patient numbers receive additional weight: Grade 1 or higher was reported for 3538 patients of 9837 patients included in placebo monotherapy cohorts that reported any anemia (36%). Among the treatment cohorts of the same publications, the reported mean frequencies of anemia grade 1 or higher were significantly higher: 35.9% (SD=32.6, range=0-100, p=0.048), compared to the placebo arms. Similar findings were observed when other cut offs were used (grade ≥2: p=0.04, grade ≥3: p=0.018), except for grade ≥4 (p=0.32).

The frequency of treatment discontinuation caused by AEs was reported in 69 of the placebo arms with a mean of 5.2% of the patients, (SD=5.2, range=0-23) and in 79 of the treatment arms with a mean of 13.5% (SD=13.9, range=0-74, p<0.0001 ANOVA). For severe adverse events (SAEs), the equivalent comparison was not significantly different between treatment and placebo arms (15.9 *vs*. 18.7%, p=0.31).

Evaluation of the influence of demographics on anemia revealed a robust finding for only one of the variables documented: Cancer diagnosis group (Figure 2, p=0.01 for grade ≥1, p=2.5×10^{-8} for grade ≥3). In contrast, there was no robust influence of median age, gender, performance status, measurable tumor, newly diagnosed *versus* relapsed/refractory, previous lines of treatments, route of treatment, total n (number of patients) in the cohort, study type (phase 1/2/3), or year of publication.

Considering the minimum number of cohorts to be averaged as four, disease-specific benchmarks can be provided for six diagnoses among the placebo monotherapy reports: Ovarian cancer (n=8 of 8 cohorts provided data, grade ≥1: anemia: 8.7%, SD=10.0, grade ≥3: 1.1% SD=1.0), colorectal (n=6 of 10, grade ≥1: 33.1% SD=17.8, grade ≥3: 4.4%, SD=3.6), hemato-oncology (n=5 of 10, grade ≥1: 8.6%, SD=7.2, grade ≥3: 0.54% SD=0.45), breast (n=4 of 8, grade ≥1: 55.6%, SD=50.3, grade ≥3: 9.1% SD=10.1), gastric (n=4 of 4, grade ≥1: 18.1%, SD=17.8, grade ≥3: 5.0%, SD=3.8) and liver cancer (n=4 of 11, grade ≥1: 25.4%, SD=22.2, grade ≥3: 4.8% SD=3.7). Among the treatment cohorts, the means were higher, however the ranking of the cancer organs was similar.

The correlation between the reported frequency of anemia and other adverse events was also examined. Among placebo monotherapy cohorts, there were robust correlations (Pearson) between anemia (grade ≥1) and thrombocytopenia (n=26, R=0.673, p=0.0002), and neutropenia (n=6, R=0.97, p=0.001), which were also confirmed when other cut-off for anemia grades were used. There were trends for positive correlations also between the frequency of anemia, and the frequencies of dyspnea, insomnia, headache, and dizziness, while the data were insufficient to analyze for correlation with tachycardia, myocardial infarct, or stroke.

Discussion

The rate of anemia grade ≥1 was reported in 46 of 100 placebo arms of randomized oncology trials; with a mean of 23.4 %. This value was dependent on the indication; it was 56% for breast cancer trial, 33% for colorectal, 25% for liver, 18% for gastric and 9% for both ovarian cancer and hemato-oncology.

A previous similar project focused on headache in AE reporting. It showed more influencing demographic variables; and among those was a counterintuitive lower frequency of headache reported among the oldest age group, and among patients with lower performance status (3). Anemia differs as it is not dependent on patient reporting but instead defined by objective laboratory values. This could explain the absence of the counterintuitive relation to the performance status in the anemia analysis.
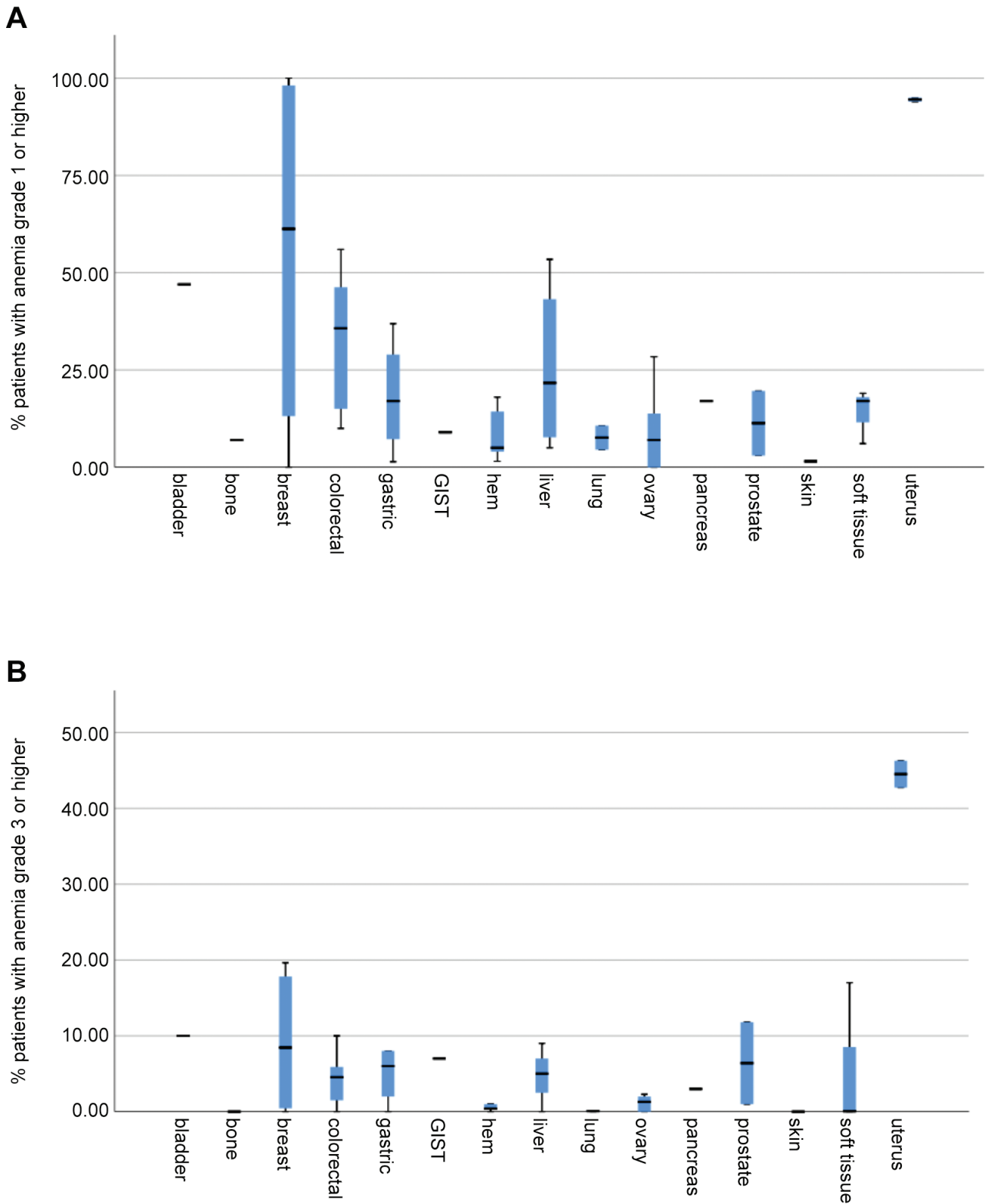
Figure 2. *Frequency of anemia grade ≥1 (A) and grade ≥3 (B) reported among placebo monotherapy arms of randomized cancer trials. The differences were significant (p=0.01 for grade ≥1, p=2.5×10⁻⁸ for grade ≥3) GIST: Gastrointestinal stroma tumor; Hem: haemato-oncology including leukemia and lymphoma.*

The pattern of anemia reporting among different indications does not match medical experience: At diagnosis, anemia is more common in leukemia patients than in liver cancer, but it was reversed here in TEAE data. The discrepancy can be explained by the definition of TEAE: Anemia present at treatment start is an adverse event (AE), but it is not a treatment emergent adverse event (TEAE). For creating bench-marks as external controls for clinical trials that lack a prospective control arm, TEAEs as reported here are the more appropriate number.

The correlation of anemia with thrombocytopenia and neutropenia indicates that most anemias reported in oncology trials reflect bone marrow related mechanisms rather than hemolysis or iron deficiency, and matches general oncology experience. Finding this correlation among clinical trials data encourages expanding the method. On single patient data levels, correlations of AEs to each other might provide a future systematic detection method for mechanisms of toxicity of novel drugs. On aggregate data level, these correlation might allow for more complete imputation of missing values.

Limitations of the study include all those associated with meta-analyses. In addition, this specific analysis combining adverse event tables has a further limitation caused by the lower limit of reporting, which is often set as 5 or 10% of patients. In consequence, adverse events of low frequency are not reported, and the means of the reported values are higher than the overall mean. Therefore, the calculated means of reported values (here: 23%) needs to be considered in conjunction with the number of cohorts in which the data were provided at all (here: 46 of 100).

The findings reported here support the novel approach of utilizing external controls to interpret AE data of single arm studies by providing bench marks and details for one specific AE: anemia. For studies including all cancer diagnoses, 23% anemia is an appropriate bench-mark, above which drug related anemia can be suspected. For specific cancer diagnoses the number should be based upon placebo arms of the same diagnosis. Some of those were generated here; for other indications, an expanded search is recommended. Model based imputation of missing values will increase the yield of such searches.

Conflicts of Interest

JW is an employee of AbbVie pharmaceuticals Inc, and owns stock. However, this project was not part of the employment, and the data interpretation reflect the authors personal opinion, not the company.

Authors' Contributions

JW: Concept, data collections, analysis, first draft, figures, revisions, finishing communications.

References

1 Ruiz-Schutz VC, Gomes LM, Mariano RC, de Almeida DVP, Pimenta JM, Dal Molin GZ, Kater FR, Yamamura R, Correa Neto NF, Maluf FC and Schutz FA: Risk of fatigue and anemia in patients with advanced cancer treated with olaparib: A meta-analysis of randomized controlled trials. Crit Rev Oncol Hematol *141*: 163-173, 2019. PMID: 31306986. DOI: 10.1016/ j.critrevonc.2019.06.012

2 Chacón MR, Enrico DH, Burton J, Waisberg FD and Videla VM: Incidence of placebo adverse events in randomized clinical trials of targeted and immunotherapy cancer drugs in the adjuvant setting: a systematic review and meta-analysis. JAMA Netw Open *1(8)*: e185617, 2018. PMID: 30646278. DOI: 10.1001/ jamanetworkopen.2018.5617

3 Wolff BJ, Wolff MJ and Wolff JE: Headache reporting in oncology trials depends on the demographics of the study population. In Vivo *35(4)*: 1939-1943, 2021. PMID: 34182466. DOI: 10.21873/invivo.12460

4 Hill AB: The environment and disease: Association or causation? Proc R Soc Med *58*: 295-300, 1965. PMID: 14283879.

5 Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, Janecek E, Domecq C and Greenblatt DJ: A method for estimating the probability of adverse drug reactions. Clin Pharmacol Ther *30(2)*: 239-245, 1981. PMID: 7249508. DOI: 10.1038/clpt.1981.154

6 Trotti A, Colevas AD, Setser A, Rusch V, Jaques D, Budach V, Langer C, Murphy B, Cumberlin R, Coleman CN and Rubin P: CTCAE v3.0: development of a comprehensive grading system for the adverse effects of cancer treatment. Semin Radiat Oncol *13(3)*: 176-181, 2003. PMID: 12903007. DOI: 10.1016/S1053-4296(03)00031-6

7 Miller TP, Li Y, Kavcic M, Troxel AB, Huang YS, Sung L, Alonzo TA, Gerbing R, Hall M, Daves MH, Horton TM, Pulsipher MA, Pollard JA, Bagatell R, Seif AE, Fisher BT, Luger S, Gamis AS, Adamson PC and Aplenc R: Accuracy of adverse event ascertainment in clinical trials for pediatric acute myeloid leukemia. J Clin Oncol *34(13)*: 1537-1543, 2016. PMID: 26884558. DOI: 10.1200/JCO.2015.65.5860

8 Zhang S, Chen Q and Wang Q: The use of and adherence to CTCAE v3.0 in cancer clinical trial publications. Oncotarget *7(40)*: 65577-65588, 2016. PMID: 27564109. DOI: 10.18632/ oncotarget.11576

9 Valle JW, Borbath I, Rosbrook B, Fernandez K and Raymond E: Sunitinib in patients with pancreatic neuroendocrine tumors: update of safety data. Future Oncol *15(11)*: 1219-1230, 2019. PMID: 30701988. DOI: 10.2217/fon-2018-0882