



Statistical reporting and table construction guidelines for publication in the *Canadian Journal of Anesthesia*

Philip M. Jones, MD, MSc · Gregory L. Bryson, MD, MSc ·
Steven B. Backman, MDCM, PhD · Sangeeta Mehta, MD ·
Hilary P. Grocott, MD

Received: 18 October 2017 / Accepted: 19 November 2017 / Published online: 5 December 2017
© Canadian Anesthesiologists' Society 2017

When publishing scientific information, data must be presented accurately, consistently, and transparently for maximally effective communication. At the *Canadian Journal of Anesthesia*, we commonly see problems in the reporting of results in the text and tables of manuscripts submitted for publication. Some of the common problems we encounter include:

- Effect sizes, such as relative risks (for binary variables) or risk differences (for binary or continuous variables) are not presented.
- The uncertainty of effect sizes is not presented with a 95% confidence interval (CI).
- Information (e.g., weight or height) is presented in either an overly precise (and clinically irrelevant) manner, by using two (or more) decimal places, or a falsely precise manner, by using decimal places that go beyond the precision of the initial measurement.
- *P* values are reported using too many decimal places.
- *P* values are reported in an imprecise manner ($P < 0.05$ or $P = NS$).
- The specific statistical test(s) that pertain to the primary (and secondary) outcome are not reported.

P. M. Jones, MD, MSc
Departments of Anesthesia & Perioperative Medicine and
Epidemiology & Biostatistics, University of Western Ontario,
London, ON, Canada

P. M. Jones, MD, MSc (✉)
Rm C3-110, University Hospital - London Health Sciences
Centre, 339 Windermere Rd, London, ON N6A 5A5, Canada
e-mail: pjones8@uwo.ca

G. L. Bryson, MD, MSc
Department of Anesthesiology and Pain Medicine, University of
Ottawa, Ottawa, ON, Canada

G. L. Bryson, MD, MSc
Clinical Epidemiology Program, Ottawa Hospital Research
Institute, Ottawa, ON, Canada

S. B. Backman, MDCM, PhD
MUHC Department of Anesthesia, McGill University Faculty of
Medicine, Montreal, QC, Canada

S. Mehta, MD
Department of Medicine, Sinai Health System, Toronto, ON,
Canada

S. Mehta, MD
Interdepartmental Division of Critical Care Medicine, University
of Toronto, Toronto, ON, Canada

H. P. Grocott, MD
Departments of Anesthesia and Surgery, University of Manitoba,
Winnipeg, MB, Canada

To address these issues, we herein provide some general guidance for statistical reporting in manuscripts submitted to the *Canadian Journal of Anesthesia* (subsequently referred to as the *Journal*). We provide example tables using fictitious data to represent the style expected for publication in the *Journal*. These examples are not meant as the final word in statistical reporting or table construction, but rather as a guide to help authors report their research. They exemplify examples of good reporting and table construction, and they highlight the means to address the most common problems encountered by the editorial staff at the *Journal*. The Senior Editorial Board developed the following guide while adhering to best practices as included in the *Scientific Style and Format*, 8th Edition by the Council of Science Editors.¹

While we recognize that *no example will fit every circumstance*, we trust that authors will endeavour to make their statistical reporting and tables conform to the provided examples as closely as possible, within the

confines of the type of data they are presenting and their study design. These examples provide a stylistic guide for authors. Furthermore, they emphasize many important statistical and reporting issues that are highly relevant to the *Journal*. Before submitting manuscripts to the *Journal*, authors should review the following guidelines and make modifications as required.

General principles

Precision

Include only the precision that is reasonable for the data being presented. For example, presenting height to one or two decimal places (e.g., 178.1 cm or 178.12 cm) is unlikely to be more informative than presenting height to the integer level (e.g., 178 cm). Similarly, unless the percentage is < 1%, it is rarely useful to report percentages to more than the integer level. Therefore, in most cases (and certainly if the sample size is < 200), report “38%”, not “38.2%”. In general, limit the reporting of measurements to the number of decimal places of precision used by the actual test. For example, do not report serum potassium measurements to two decimal places when the laboratory measure is capable of only one decimal place.

Measures of effect size

The expectation is that data are first summarized for each group and then presented with the effect size, 95% CI of the effect size, and (optionally) a *P* value. This format is expected in both the text and the tables. The only difference between reporting data in the text *vs* in the tables is that separate columns are used for the required information when using tables.

For relative risks, hazard ratios, and odds ratios (OR), use two decimal places for numbers < 5, and use one decimal place when ≥ 5 (e.g., OR, 1.47; 95% CI, 0.14 to 7.1). Note that, in general, absolute risk differences are typically more clinically meaningful than relative risks since, for rare events, a relative risk may not reflect clinical importance.² For example, a mortality incidence of 0.1% in an intervention group compared with 0.2% in a control group has a risk difference of 0.1% (and a number needed to treat of 1,000) but a relative risk of 0.5.

For differences in means or medians, use a level of precision appropriate for the outcome. For example, presenting the difference in time to extubation (in hours) to two decimal places is unlikely to be clinically useful, but one decimal place (which is essentially six minutes of

precision) is reasonable. This is shown in Table 1b (“time to first tracheal extubation”).

Examples (as would be reported in the Abstract or body of the manuscript)

Continuous variables: “The mean (SD) visual analogue scale pain scores were 5.6 (3.4) and 3.3 (2.9) in the placebo and ibuprofen groups, respectively (difference in means, -2.3 ; 95% CI, -1.2 to -3.4 ; $P = 0.03$).”

Example for binary variables (preferred, using absolute risk difference):

“The incidence of hemodialysis was 4/200 (2%) in the intervention group and 10/200 (5%) in the control group (risk difference, -3% ; 95% CI, -6.6 to 0.6% ; $P = 0.10$).”

Example for binary variables (acceptable but not preferred, using relative risk):

“The incidence of hemodialysis was 4/200 (2%) in the intervention group and 10/200 (5%) in the control group (relative risk, 0.4; 95% CI, 0.13 to 1.25; $P = 0.10$).”

Hypothesis tests

It is common for authors to state that, for example, “Fisher’s exact test or the Chi square test was used, *as appropriate*” or “a Student’s *t* test or Wilcoxon rank-sum test was used, *as appropriate*”. In the Methods section, it is important to state the exact conditions that led to the use of a particular hypothesis test (i.e., stating “*as appropriate*” does not provide the reader with enough information to understand why a particular test was used). Additionally, in the tables, ensure that the hypothesis test used for each variable is indicated clearly using footnotes (see Table 1b).

P values

Do not use the term “not significant” or “NS” in tables – instead give actual *P* values. *P* values should be limited to two decimal places, unless $P < 0.01$ (in which case three decimal places should be used). If the *P* value is less than 0.001, do not include the calculated *P* value – instead present it as “ $P < 0.001$ ”. An exception to this is genomic research where, due to the multiplicity concerns inherent in this type of research, it is routine to present actual *P* values, even if they are extremely small.

Interquartile ranges and confidence intervals

Since negative values rarely occur in an interquartile range, the values are presented with a dash between the 25th and 75th centiles, and the values are enclosed within square brackets (e.g., [61-76]). In contrast, the confidence limits in 95% confidence intervals commonly comprise negative

Table 1a Characteristics of the patients at baseline in an RCT

Characteristic	Levosimendan N = 231	Placebo ^a N = 233
Female, <i>n</i> /total <i>N</i> (%)	46/231 (20%)	51/232 (22%)
Age (yr), mean (SD)	66.1 (8.8)	65.8 (9.2)
Height (cm), mean (SD)	171 (10.4)	171 (8.4)
	<i>n</i> = 230	<i>n</i> = 231
Serum creatinine ($\mu\text{mol}\cdot\text{L}^{-1}$), median [IQR]	85 [55-143]	84 [51-140]
Left ventricular grade, <i>n</i> /total <i>N</i> (%)		
1 (LVEF > 54%)	153/230 (67%)	162/233 (70%)
2 (LVEF 40 - 54%)	57/230 (25%)	51/233 (22%)
3 (LVEF 20 - 39%)	18/230 (8%)	19/233 (8%)
4 (LVEF < 20%)	2/230 (0.9%)	1/233 (0.4%)
Preoperative Medications		
Beta-blocker, <i>n</i> /total <i>N</i> (%)	162/231 (70%)	168/233 (72%)
ACE inhibitor or angiotensin II receptor blocker, <i>n</i> /total <i>N</i> (%)	155/228 (68%)	154/233 (66%)
ASA, <i>n</i> /total <i>N</i> (%)	178/230 (77%)	177/233 (76%)

^aThe placebo was identical in colour to the levosimendan

ACE = angiotensin-converting enzyme; ASA = acetylsalicylic acid; IQR = interquartile range; LVEF = left ventricular ejection fraction; RCT = randomized clinical trial; SD = standard deviation

Denominators that do not equal the sample sizes are due to missing data. If no denominator is given, all data were present

Table 1b Outcomes from an RCT

Outcome	Levosimendan N = 231	Placebo N = 233	Relative Risk (95% CI)	<i>P</i> value
<i>Primary Outcome</i>				
Composite of prolonged ICU stay (≥ 48 hr) and 30-day all-cause mortality, <i>n</i> /total <i>N</i> (%)	57/231 (25%)	70/233 (30%)	0.82 (0.61 to 1.11)	0.21 ^a
<i>Categorical Secondary Outcomes</i>				
Prolonged ICU stay (≥ 48 hr), <i>n</i> /total <i>N</i> (%)	54/231 (23%)	70/233 (30%)	0.78 (0.57 to 1.06)	< 0.001 ^a
30-day all-cause mortality, <i>n</i> /total <i>N</i> (%)	6/231 (2.6%)	3/233 (1.3%)	2.0 (0.51 to 8.0)	0.34 ^a
<i>Continuous Secondary Outcomes</i>				
			<i>Difference in Means (95% CI)</i>	
Troponin T 6 hr after ICU admission ($\text{ng}\cdot\text{L}^{-1}$), mean (SD)	642 (809) <i>n</i> = 222	576 (620) <i>n</i> = 228	66 (−67 to 199)	0.33 ^b
Peak postoperative serum creatinine ($\mu\text{mol}\cdot\text{L}^{-1}$), mean (SD)	99 (36) <i>n</i> = 231	97 (44) <i>n</i> = 233	2 (−6 to 9)	0.62 ^b
			<i>Difference in Medians (95% CI)</i>	
Time to first tracheal extubation (hr), median [IQR]	5.2 [3.6-9.3] <i>n</i> = 229	5.3 [3.7-8.3] <i>n</i> = 231	−0.1 (−1.1 to 0.9)	0.95 ^c
Time to discharge from ICU (days), median [IQR]	1.10 [0.93-1.91] <i>n</i> = 229	1.13 [0.97-2.09] <i>n</i> = 231	−0.03 (−0.18 to 0.12)	0.17 ^c
Time to discharge from hospital (days), median [IQR]	5.9 [5.0-8.0] <i>n</i> = 225	5.9 [4.9-8.0] <i>n</i> = 230	−0.07 (−0.37 to 0.22)	0.86 ^c

Relative risks are for levosimendan relative to placebo; differences are (levosimendan — placebo). Denominators that do not equal sample sizes are due to either missing data or patients dying before the outcome could occur

^aFisher's exact test

^bStudent's *t* test

^c95% confidence interval from 0.5 quantile (median) regression with bootstrapped standard errors (10,000 replications); *P* value from log-rank test

CI = confidence interval; ICU = intensive care unit; IQR = interquartile range; RCT = randomized clinical trial; SD = standard deviation

Table 2a Characteristics of the patients at baseline in an observational study

Characteristic	Control Group <i>N</i> = 307,100	Intervention Group <i>N</i> = 5,900	<i>P</i> value	Standardized Differences in Unweighted Sample (%)	Standardized Differences in Weighted Sample (%)
Female	171,397 (55.8%)	2,764 (46.5%)	< 0.001	17.9	2.9
Age (yr), mean (SD)	59.8 (15.6)	59.9 (15.7)	0.47	0.4	3.3
Comorbidities					
Charlson comorbidity index					
Mean (SD)	0.86 (1.49)	0.97 (1.63)	< 0.001	7.3	6.3
Median [IQR]	0 [0-2]	0 [0-2]	-	-	-
Hypertension	161,498 (53.3%) <i>N</i> = 303,000	3,234 (55.7%) <i>N</i> = 5,803	0.005	3.6	-2.5
Coronary artery disease	92,959 (30.3%)	1,676 (28.2%)	0.001	-4.7	1.7

All numbers are *n*/group *N* (%) unless otherwise specified. Comorbidities included a five-year look-back window. *P* values are from the Chi square test or Student’s *t* test. The Charlson comorbidity index is a list of 17 comorbidities identified by International Statistical Classification of Diseases-10 (ICD-10) codes, each of which is assigned a weight from 1–6. The score starts at 0 for healthy patients (no comorbidities identified). Higher scores indicate the presence of more comorbidities. IQR = interquartile range; SD = standard deviation

Table 2b Outcomes from an observational study

Outcome	Control Group <i>N</i> = 307,100	Intervention Group <i>N</i> = 5,900	Unadjusted Risk Difference (95% CI) <i>N</i> = 313,000	Adjusted Risk Difference (95% CI) <i>N</i> = 313,000
Primary outcome of all-cause death, hospital readmission, or major complication within 30 days — <i>n</i> / <i>N</i> (%)	90,000 (29.3%) <i>N</i> = 307,000	2,580 (43.7%) <i>N</i> = 5,800	14.4% (12.7 to 15.4) <i>P</i> < 0.001 <i>N</i> = 312,800	6.6% (4.6 to 9.2) <i>P</i> < 0.001 <i>N</i> = 312,800
Duration of acute hospital stay (days), mean (SD)	7.4 (13.2)	12.2 (18.1)	4.8 days (4.5 to 5.2) <i>P</i> < 0.001	1.1 days (0.70 to 1.68) <i>P</i> < 0.001
Emergency Department visits within 90 days of the index surgery, mean (SD)	0.53 (1.27)	0.62 (1.48)	0.09 visits (0.06 to 0.12) <i>P</i> < 0.001	0.04 visits (-0.04 to 0.10) <i>P</i> = 0.36

Adjusted risk differences obtained from inverse probability of exposure weighting based on propensity scores. The propensity score was estimated using multivariable logistic regression with receipt of the intervention as the dependent variable and a vector of covariates decided upon *a priori* as the independent variables (sex, age, comorbidities with a five-year look-back window [i.e., hypertension, coronary artery disease, congestive heart failure, peripheral vascular disease, diabetes, previous stroke or transient ischemic attack, chronic liver disease, cancer, chronic renal disease, and chronic obstructive pulmonary disease], and the type of surgery performed). Risk differences are for the intervention group relative to the control group. CI = confidence interval; SD = standard deviation

values. To avoid the ambiguity of having a dash followed by a negative sign (i.e., “- -”), the values are presented with the word “to” between the confidence limits. The values are enclosed within parentheses [e.g., (-56.8 to -1.55)].

Redundancy

In general, data should be presented in either a figure or a table, but not in both. In particular, authors should try to avoid lengthy and cumbersome descriptions of findings in the Results section of the manuscript. With many numbers,

the text can be difficult to follow, and the information can be more easily conveyed in a well-constructed table or figure. Clearly explain the findings and refer the reader to the table(s) or figure(s).

Encapsulation

Table captions should fully explain the contents of a table, including all abbreviations, and they should briefly explain clinical scoring systems that may be unfamiliar to readers (such as the Charlson comorbidity index — see Table 2a). For tables presenting regression coefficients, all analyzed

covariates must be presented, not just those that were statistically significant, and their 95% CIs must be included. For tables presenting results analyzed using regression techniques, the covariates included in the regression model must be listed in the table's caption, even if they are already present in the main manuscript.

Tables for reports of randomized clinical trials

Table 1a presents demographic variables in a randomized clinical trial (RCT). We have limited the rows to present each type of commonly used data: binary, multilevel categorical, continuous (normally distributed), and continuous (non-normally distributed). Also, notice the use of footnotes using superscripted, consecutive lowercase letters (not daggers, asterisks, or other symbols) and the proper definitions for abbreviations.

Note that no P values are included for pre-randomization demographic variables when reporting an RCT.³

It is acceptable to exclude denominators for binary or categorical variables only if no data are missing. If data are missing, please include the overall number of participants with available data by including the denominator (as illustrated in this example). This also applies to continuous variables (as in height) where missing data for several participants are explicitly acknowledged in the table.

Note the indentation of rows that fall within a global category (as in left ventricular grade).

Table 1b presents the results of a randomized clinical trial. The key point for this table is that the results for each group are displayed first, followed by the effect size (and the 95% CI for the effect size). The effect size for every type of data should be quantified irrespective of being binary/categorical data or continuous variables. Although it is straightforward to calculate the effect size and confidence interval for a difference in means, it is more complex to calculate these for a difference in medians. Various choices are available, including use of the Hodges-Lehmann method or bootstrapping.⁴ Consultation with a statistician is strongly encouraged.

Tables for reports of observational studies

Table 2a presents demographic variables from an observational study. This study used inverse probability of exposure *weighting* using the propensity score for analysis, so the column heading indicates “unweighted” and “weighted” samples. For studies analyzed using propensity score *matching*, these columns would be slightly renamed to indicate the samples “before

matching” and “after matching”. Note that it is acceptable to include P values here since the study is not an RCT. In addition, variables with missing data are explicitly mentioned and the missing data are not concealed from the reader. The caption for Table 2a not only defines the abbreviations used but also gives a brief description of the Charlson comorbidity index, which informs readers who are unfamiliar with it.

Table 2b presents the outcomes from an observational study. Note that the unadjusted results are presented in the column adjacent to the adjusted results and that missing data are explicitly mentioned. Effect sizes are given along with the confidence intervals around the effect size. The caption for Table 2b includes the covariates used in the analytical model to encapsulate the results presented.

Directives pour la communication des statistiques et l'élaboration de tableaux pour publication dans le *Journal canadien d'anesthésie*

Lors de la publication de renseignements scientifiques, il faut s'efforcer de bien communiquer les données. Pour ce faire, celles-ci doivent être présentées de façon précise, cohérente et transparente. Au *Journal canadien d'anesthésie*, nous sommes souvent confrontés à des problèmes de communication des résultats dans les textes et tableaux des manuscrits soumis pour publication. Voici une liste, non exhaustive, des problèmes fréquemment rencontrés :

- Les ampleurs d'effet, tels que les risques relatifs (pour les variables binaires) ou les différences de risque (pour les variables binaires ou continues) ne sont pas présentées.
- L'incertitude des ampleurs d'effet n'est pas présentée avec un intervalle de confiance (IC) de 95 %.
- Certaines informations (par ex., poids ou taille) sont présentées de façon excessivement précise (et non pertinente d'un point de vue clinique), en utilisant des chiffres à deux décimales (ou plus), ou d'une manière faussement précise, en utilisant des décimales qui vont au-delà de la précision de la mesure initiale.
- Les valeurs P sont rapportées avec trop de décimales.
- Les valeurs P sont rapportées de façon imprécise ($P < 0,05$ ou $P = \text{NS}$).
- Le(s) test(s) statistique(s) spécifique(s) appliqué(s) au critère d'évaluation primaire (ou aux critères secondaires) n'est/ne sont pas rapporté(s).

Afin de pallier à ces problèmes, nous présentons ici des directives générales concernant la communication des statistiques dans les manuscrits soumis au *Journal canadien d'anesthésie* (le *Journal*). Nous proposons au lecteur des tableaux d'exemples utilisant des données fictives afin de représenter le style attendu pour publication dans le *Journal*. Ces exemples ne constituent en rien la seule façon de communiquer des résultats statistiques ou d'élaborer des tableaux; plutôt, leur fonction est de guider les auteurs qui désirent publier leurs recherches. Ces exemples montrent ce que nous entendons par la bonne communication des résultats et l'élaboration correcte de tableaux, et ils proposent diverses façons de corriger les problèmes les plus fréquemment rencontrés par l'équipe rédactionnelle du *Journal*. Le Comité de rédaction principal a mis au point le guide suivant en respectant les meilleures pratiques telles qu'elles sont présentées dans le guide anglophone *Scientific Style and Format*, 8^{ème} édition, publié par le *Council of Science Editors*.¹

Nous sommes tout à fait conscients qu'*aucun exemple ne s'appliquera à toutes les situations*; cependant, nous sommes confiants que les auteurs s'efforceront de rendre la présentation de leurs statistiques et tableaux conforme aux exemples illustrés autant que possible – dans les limites du type de données et de la méthodologie de leur étude. Ces exemples constituent un guide de style destiné aux auteurs. En outre, ils soulignent plusieurs questions importantes en matière de statistiques et de communication des résultats tout à fait pertinentes au *Journal*. Avant de soumettre un manuscrit au *Journal*, nous encourageons les auteurs à examiner les directives suivantes et à apporter les modifications nécessaires à leur manuscrit, le cas échéant.

Principes généraux

Précision

N'inclure que le degré de précision raisonnable pour les données présentées. Par exemple, le fait de présenter la taille d'un patient avec une ou deux décimales (par ex., 178,1 cm ou 178,12 cm) ne fournira probablement pas plus d'informations que de la présenter sous forme d'entier relatif (par ex., 178 cm). De la même manière, il est rarement utile de rapporter les pourcentages de façon plus précise que l'entier relatif, sauf si le pourcentage est < 1 %. Ainsi, dans la plupart des cas (et certainement si la taille d'échantillon est < 200), rapporter « 38 % » et non « 38,2 % ». En règle générale, limiter la communication des mesures au nombre de décimales de précision du test réalisé. Par exemple, ne pas communiquer les mesures de potassium sérique avec deux décimales si la mesure de laboratoire n'affiche qu'une décimale.

Mesures de l'ampleur de l'effet

On s'attend à ce que les données soient d'abord résumées pour chaque groupe, puis présentées avec l'ampleur de l'effet, l'IC 95 % de l'ampleur de l'effet, et (peut-être) une valeur *P*. Ce format devrait être la norme tant dans le corps du texte que dans les tableaux. La seule différence de communication des données entre le corps du texte et les tableaux est l'utilisation de colonnes distinctes pour présenter les renseignements dans les tableaux.

En ce qui a trait aux risques relatifs, rapports de risques et rapports de cotes (RC - *odds ratio*), utiliser des chiffres à 2 décimales pour les nombres < 5, et des chiffres à une décimale pour les nombres ≥ 5 (par ex., RC, 1,47; IC 95 %, 0,14 à 7,1). Prendre note qu'en règle générale, les différences de risque absolu sont souvent plus importantes d'un point de vue clinique que les risques relatifs étant donné que, en cas d'événement rare, un risque relatif pourrait ne pas refléter son importance clinique.² Par exemple, si l'incidence de mortalité est de 0,1 % dans le groupe intervention, comparée à 0,2 % dans le groupe témoin, la différence de risque est de 0,1 % (et le nombre de sujets à traiter est de 1000), mais le risque relatif est de 0,5.

En ce qui a trait aux différences de moyennes ou de médianes, utiliser un degré de précision adapté au résultat. Par exemple, il est probablement peu utile, d'un point de vue clinique, de présenter la différence de temps jusqu'à l'extubation (en heures) avec deux décimales, mais un chiffre à une décimale (ce qui correspond à une précision de six minutes) est raisonnable. Le Tableau 1b (« Temps jusqu'à la première extubation trachéale ») illustre cet exemple.

Exemples (tels qu'ils seraient rapportés dans le Résumé ou dans le corps du texte du manuscrit)

Variables continues : « Les scores de douleur moyens (ÉT) sur une échelle visuelle analogique étaient de 5,6 (3,4) et 3,3 (2,9) dans les groupes placebo et ibuprofène, respectivement (différence de moyennes, -2,3; IC 95 %, -1,2 à -3,4; *P* = 0,03). »

Exemple pour les variables binaires (présentation préférée, avec la différence de risque absolu) :

« L'incidence d'hémodialyse était de 4/200 (2 %) dans le groupe intervention et de 10/200 (5 %) dans le groupe témoin (différence de risque, -3 %; IC 95 %, -6,6 à 0,6 %; *P* = 0,10). »

Exemple pour les variables binaires (présentation acceptable mais non préférée, avec le risque relatif) :

« L'incidence d'hémodialyse était de 4/200 (2 %) dans le groupe intervention et de 10/200 (5 %) dans le groupe

témoin (risque relatif, 0,4; IC 95 %, 0,13 à 1,25; $P = 0,10$). »

Tests d'hypothèse

Il arrive fréquemment que les auteurs déclarent, par exemple : « Le test exact de Fisher ou le test du Chi carré a été utilisé, *le cas échéant* », ou encore « un test t de Student ou un test de Wilcoxon a été utilisé, *tel qu'approprié* ». Dans la section Méthode, il est important de décrire les conditions exactes ayant mené à l'utilisation d'un test d'hypothèse en particulier (en d'autres mots, le fait d'écrire « tel qu'approprié » ou « le cas échéant » ne fournit pas suffisamment de renseignements au lecteur pour qu'il puisse comprendre pourquoi un test en particulier a été utilisé). En outre, dans les tableaux, il faut s'assurer que le test d'hypothèse utilisé pour chaque variable soit clairement indiqué dans les notes explicatives (voir Tableau 1b).

Valeurs P

Ne pas utiliser le terme « non significatif » ou « NS » dans les tableaux – indiquer les valeurs P réelles à la place. Les valeurs P devraient se limiter à des chiffres à deux décimales, à moins que $P < 0,01$ (dans quel cas il faut utiliser trois décimales). Si la valeur P est inférieure à 0,001, ne pas inclure la valeur P calculée – présenter plutôt

cette valeur ainsi : $P < 0,001$. La recherche génomique fait exception à cette règle; en effet, en raison des questions de multiplicité inhérentes à ce type de recherche, il est habituel de présenter les valeurs P réelles, même si elles sont extrêmement petites.

Écarts interquartiles et intervalles de confiance

Les valeurs négatives étant rares dans un écart interquartile, les valeurs sont présentées avec un tiret et entre crochets entre le 25^e et le 75^e percentiles (par ex., [61-76]). En revanche, les limites de confiance dans les intervalles de confiance de 95 % comprennent fréquemment des valeurs négatives. Afin d'éviter toute ambiguïté soulevée par l'écriture consécutive d'un tiret et d'un signe négatif (c.-à-d., « - - »), le mot « à » séparera les valeurs entre les limites de confiance. Les valeurs sont présentées entre parenthèses [par ex., (-56,8 à -1,55)].

Redondance

Règle générale, les données devraient être présentées sous forme de figures ou de tableaux, mais pas les deux. En particulier, les auteurs doivent s'efforcer d'éviter les descriptions longues et lourdes de leurs résultats dans la section Résultats du manuscrit. Un texte énumérant beaucoup de chiffres sera difficile à lire, et les informations seront mieux transmises par le biais d'un

Tableau 1a Caractéristiques de base des patients dans une ERC

Caractéristique	Lévosimendan $N = 231$	Placebo ^a $N = 233$
Sexe féminin, n /total N (%)	46/231 (20 %)	51/232 (22 %)
Âge (années), moyenne (ÉT)	66,1 (8,8)	65,8 (9,2)
Taille (cm), moyenne (ÉT)	171 (10,4)	171 (8,4)
	$n = 230$	$n = 231$
Créatinine sérique ($\mu\text{Mo}\cdot\text{L}^{-1}$), médiane [ÉIQ]	85 [55-143]	84 [51-140]
Grade ventriculaire gauche, n /total N (%)		
1 (FEVG > 54 %)	153/230 (67 %)	162/233 (70 %)
2 (FEVG 40 – 54 %)	57/230 (25 %)	51/233 (22 %)
3 (FEVG 20 - 39 %)	18/230 (8 %)	19/233 (8 %)
4 (FEVG < 20 %)	2/230 (0,9 %)	1/233 (0,4 %)
Médicaments préopératoires		
Bêta-bloqueurs, n /total N (%)	162/231 (70 %)	168/233 (72 %)
Inhibiteur de l'enzyme de conversion de l'angiotensine ou bloqueur des récepteurs de l'angiotensine II, n /total N (%)	155/228 (68 %)	154/233 (66 %)
Aspirine, n /total N (%)	178/230 (77 %)	177/233 (76 %)

^aLe placebo était de couleur identique au lévosimendan

ÉIQ = écart interquartile; ERC = étude randomisée contrôlée; ÉT = écart type; FEVG = fraction d'éjection du ventricule gauche

Les dénominateurs qui ne sont pas égaux aux tailles d'échantillon sont dus à des données manquantes. Si aucun dénominateur n'est spécifié, toutes les données étaient présentes

Tableau 1b Critères d'évaluation d'une ERC

Critère d'évaluation	Lévosimendan N = 231	Placebo N = 233	Risque relatif (IC 95 %)	Valeur P
<i>Critère d'évaluation principal</i>				
Mesure composite d'un séjour prolongé à l'USI (≥ 48 h) et de mortalité à 30 jours, toutes causes confondues, n/total N(%)	57/231 (25 %)	70/233 (30 %)	0,82 (0,61 à 1,11)	0,21 ^a
<i>Critères d'évaluation secondaires catégoriels</i>				
Séjour prolongé à l'USI (≥ 48 h), n/total N (%)	54/231 (23 %)	70/233 (30 %)	0,78 (0,57 à 1,06)	< 0,001 ^a
Mortalité à 30 jours, toutes causes confondues, n/total N (%)	6/231 (2,6 %)	3/233 (1,3 %)	2,0 (0,51 à 8,0)	0,34 ^a
<i>Critères d'évaluation secondaires continus</i>				
			<i>Différence de moyennes (IC 95 %)</i>	
Troponine T 6 h après l'admission à l'USI (ng·L ⁻¹), moyenne (ÉT)	642 (809) n = 222	576 (620) n = 228	66 (-67 à 199)	0,33 ^b
Créatinine sérique postopératoire maximale (μMo·L ⁻¹), moyenne (ÉT)	99 (36) n = 231	97 (44) n = 233	2 (-6 à 9)	0,62 ^b
			<i>Différence de médianes (IC 95 %)</i>	
Délai jusqu'à la première extubation trachéale (h), médiane [ÉIQ]	5,2 [3,6-9,3] n = 229	5,3 [3,7-8,3] n = 231	-0,1 (-1,1 à 0,9)	0,95 ^c
Délai jusqu'au congé de l'USI (jours), médiane [ÉIQ]	1,10 [0,93-1,91] n = 229	1,13 [0,97-2,09] n = 231	-0,03 (-0,18 à 0,12)	0,17 ^c
Délai jusqu'au congé de l'hôpital (jours), médiane [ÉIQ]	5,9 [5,0-8,0] n = 225	5,9 [4,9-8,0] n = 229	-0,07 (-0,37 à 0,22)	0,86 ^c

Les risques relatifs concernent le lévosimendan par rapport au placebo; les différences sont (lévosimendan – placebo). Les dénominateurs qui n'égalent pas les tailles d'échantillon sont dus à des données manquantes ou à des patients décédant avant la survenue du critère d'évaluation

^aTest exact de Fisher

^bTest *t* de Student

^cIntervalle de confiance à 95 % d'une régression de quantile de 0,5 (médiane) avec des erreurs type d'auto-amorçage (10 000 répétitions); valeur *P* issue d'un test logarithmique par rangs

ÉIQ = écart interquartile; ERC = étude clinique randomisée; ÉT = écart type; IC = intervalle de confiance; USI = unité de soins intensifs

tableau ou d'une figure bien conçu. Il convient d'expliquer clairement les résultats et de référer le lecteur au(x) tableau(x) ou figure(s).

Encapsulation

Les légendes des tableaux devraient expliquer clairement et exhaustivement le contenu d'un tableau, y compris toutes les abréviations. Elles devraient également fournir une brève explication de tout système de notation que le lecteur pourrait ne pas connaître (comme par ex. l'indice de comorbidité de Charlson – voir Tableau 2a). En ce qui touche aux tableaux présentant des coefficients de régression, toutes les covariables analysées doivent être présentées, et pas uniquement les covariables statistiquement significatives, et leur IC 95 % doivent être inclus. En ce qui a trait aux tableaux présentant des résultats analysés à l'aide de techniques de régression, les covariables incluses dans le modèle de régression doivent

être énumérées dans la légende du tableau, même si elles sont déjà présentées dans le manuscrit principal.

Tableaux accompagnant les comptes rendus d'études cliniques randomisées

Le tableau 1a présente les variables démographiques d'une étude clinique randomisée (ERC). Nous avons limité les rangées pour présenter chaque type de donnée fréquemment utilisée : binaire, catégorielle à niveaux multiples, continue (distribution normale) et continue (à distribution non normale). Noter également l'utilisation de notes de bas de page en lettres minuscules consécutives insérées en position supérieure (pas de croix, astérisques ou autres symboles) et les définitions des abréviations.

Noter que, lors de la présentation d'une ERC, aucune valeur *P* n'est incluse pour les variables démographiques pré-randomisation.³

Tableau 2a Caractéristiques de base des patients dans une étude observationnelle

Caractéristique	Groupe témoin <i>N</i> = 307 100	Groupe intervention <i>N</i> = 5900	Valeur <i>P</i>	Différences normalisées dans un échantillon non pondéré (%)	Différences normalisées dans un échantillon pondéré (%)
Sexe féminin	171 397 (55,8 %)	2764 (46,5 %)	< 0,001	17,9	2,9
Âge (années), moyenne (ÉT)	59,8 (15,6)	59,9 (15,7)	0,47	0,4	3,3
Comorbidités					
Indice de comorbidité de Charlson					
moyenne (ÉT)	0,86 (1,49)	0,97 (1,63)	< 0,001	7,3	6,3
médiane [ÉIQ]	0 [0-2]	0 [0-2]	-	-	-
Hypertension	161 498 (53,3 %) <i>N</i> = 303 000	3234 (55,7 %) <i>N</i> = 5803	0,005	3,6	-2,5
Maladie coronarienne	92 959 (30,3 %)	1676 (28,2 %)	0,001	-4,7	1,7

Tous les chiffres sont *n*/groupe *N* (%) sauf indication contraire. Les comorbidités ont été évaluées en se fondant sur les antécédents des cinq dernières années. Les valeurs *P* proviennent du test du Chi carré ou du test *t* de Student. L'indice de comorbidité de Charlson est une liste de 17 comorbidités identifiées par les codes de la Classification statistique internationale des maladies (CIM-10), qui assigne un poids de 1-6 à chacune des comorbidités. Le score débute à 0 pour les patients en bonne santé (aucune comorbidité identifiée). Des scores plus élevés indiquent la présence de comorbidités. ÉIQ = écart interquartile; ÉT = écart type

Tableau 2b Critères d'évaluation d'une étude observationnelle

Critère d'évaluation	Groupe témoin <i>N</i> = 307 100	Groupe intervention <i>N</i> = 5900	Différence de risque non ajustée (IC 95 %) <i>N</i> = 313 000	Différence de risque ajustée (IC 95 %) <i>N</i> = 313 000
Critère d'évaluation principal incluant les décès, toutes causes confondues, réadmission à l'hôpital, ou complication majeure à 30 jours – <i>n</i> / <i>N</i> (%)	90 000 (29,3 %) <i>N</i> = 307 000	2580 (43,7 %) <i>N</i> = 5800	14,4 % (12,7 à 15,4) <i>P</i> < 0,001 <i>N</i> = 312 800	6,6 % (4,6 à 9,2 %) <i>P</i> < 0,001 <i>N</i> = 312 800
Durée de séjour à l'hôpital en soins aigus (jours), moyenne (ÉT)	7,4 (13,2)	12,2 (18,1)	4,8 jours (4,5 à 5,2) <i>P</i> < 0,001	1,1 jour (0,70 à 1,68) <i>P</i> < 0,001
Visites à l'urgence au cours des 90 jours de la chirurgie initiale, moyenne (ÉT)	0,53 (1,27)	0,62 (1,48)	0,09 visite (0,06 à 0,12) <i>P</i> < 0,001	0,04 visite (-0,04 à 0,10) <i>P</i> = 0,36

Les différences de risque ajustées ont été obtenues d'une pondération inverse sur la probabilité d'exposition fondée sur les scores de propension. Le score de propension a été estimé à l'aide d'une régression logistique multivariée avec le début de l'intervention en tant que variable dépendante et un vecteur de covariables déterminé *a priori* en tant que variables indépendantes (sexe, âge, comorbidités des 5 dernières années [c.-à-d. hypertension, maladie coronarienne, insuffisance cardiaque, maladie vasculaire périphérique, diabète, accident vasculaire cérébral ou ischémique cérébral transitoire, maladie hépatique chronique, cancer, néphropathie chronique, et maladie pulmonaire obstructive chronique], et type de chirurgie réalisé). Les différences de risque concernent le groupe intervention par rapport au groupe témoin. ÉT = écart type; IC = intervalle de confiance

Il est acceptable d'exclure les dénominateurs des variables binaires ou catégoriques seulement si aucune donnée ne manque. Si des données manquent, inclure le nombre global de participants avec les données disponibles en incluant le dénominateur (tel qu'illustré dans cet exemple). Cette règle s'applique également aux variables continues (telles que la taille) où il est fait mention explicite des données manquantes de plusieurs participants dans le tableau.

Noter l'indentation des rangées qui tombent sous une catégorie globale (comme par ex. grade ventriculaire gauche).

Le tableau 1b présente les résultats d'une étude clinique randomisée. Le point clé de ce tableau est que les résultats de chaque groupe sont présentés en premier, puis vient l'ampleur de l'effet (et l'IC 95 % pour l'ampleur de l'effet). L'ampleur de l'effet de chaque type de données devrait être quantifiée, qu'il s'agisse de données

binaires/catégorielles ou de variables continues. Bien qu'il soit simple de calculer l'ampleur de l'effet et l'intervalle de confiance pour une différence de moyennes, il est plus difficile de calculer ces valeurs pour une différence de médianes. Diverses options sont possibles, notamment l'utilisation d'une méthode de Hodges-Lehmann ou d'auto-amorçage (*bootstrapping*).⁴ Il est fortement recommandé de consulter un statisticien.

Tableaux accompagnant les comptes rendus d'études observationnelles

Le tableau 2a présente les variables démographiques d'une étude observationnelle. Cette étude a utilisé une pondération inverse sur la probabilité d'exposition à l'aide d'un score de propension pour l'analyse, c'est pourquoi l'entête de la colonne indique des échantillons « non pondérés » et « pondérés ». Pour les études analysées à l'aide d'une méthode d'appariement des scores de propension, ces colonnes devraient être légèrement renommées pour indiquer les échantillons « avant appariement » et « après appariement ». Noter qu'il est acceptable d'inclure des valeurs *P* ici étant donné que l'étude n'est pas une ERC. En outre, les variables comprenant des données manquantes sont mentionnées de façon explicite et les données manquantes ne sont pas dissimulées au lecteur. La légende du tableau 2a définit non seulement les abréviations utilisées, mais décrit également brièvement l'indice de comorbidité de Charlson, pour l'expliquer au lecteur qui ne le connaîtrait pas.

Le tableau 2b présente les résultats d'une étude observationnelle. Noter que les résultats non ajustés sont présentés dans la colonne adjacente aux résultats ajustés, et que les données manquantes sont mentionnées de façon explicite. Les ampleurs d'effet sont présentées avec les intervalles de confiance autour de l'ampleur de l'effet. La légende du tableau 2b inclut les covariables utilisées dans le modèle analytique afin d'encapsuler les résultats présentés.

Conflicts of interest None declared.

Editorial responsibility This submission was handled by Dr. A. Stéphane Lambert, Associate Editor, *Canadian Journal of Anesthesia*.

Conflit d'intérêt Aucun.

Responsabilité éditoriale Cet article a été traité par Dr A. Stéphane Lambert, rédacteur adjoint, *Journal canadien d'anesthésie*.

References

1. Council of Science Editors. Scientific Style and Format – The CSE Manual for Authors, Editors, and Publishers – 8th Edition. ISBN: 978-0226116495. Chicago: University of Chicago Press; 2014.
2. Nuovo J, Melnikow J, Chang D. Reporting number needed to treat and absolute risk reduction in randomized controlled trials. *JAMA* 2002; 287: 2813-4.
3. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000; 355: 1064-9.
4. Haukoos JS, Lewis RJ. Advanced statistics: bootstrapping confidence intervals for statistics with “difficult” distributions. *Acad Emerg Med* 2005; 12: 360-5.