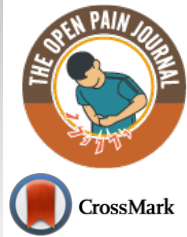




The Open Pain Journal

Content list available at: <https://openpainjournal.com>



RESEARCH ARTICLE

Sample Size Estimation for a Non-inferiority Pain Management Trial

Anadya Prakash Tripathi^{1,*} and Rama Shanker¹

¹Department of Statistics, Assam University, Silchar, Assam, India

Abstract:

Introduction:

Measuring pain and pain relief are the primary concerns in pain management. Sample size estimation in pain management with non-inferiority (NI) study design and assessment of specific-NI margin endpoints may be challenging as pain and its improvement are measured and reported on different endpoints.

Methods:

Multiple endpoints were reported frequently to measure pain and pain improvement. The sum of pain intensity difference (SPID(0-t)) at a specific time is the recommended endpoint for the measurement of pain by the United States Food and Drug Administration. Statistical information on SPID and other endpoints reported in multiple works in the literature (preferably from placebo-controlled trials) was collected and compared to identify a suitable NI margin. A difference of 20% was considered the default NI margin for evaluation, and the sample size was calculated for each endpoint.

Results:

The sample size based on the FDA-recommended primary endpoint SPID was found to be larger. This may be a concern for overall clinical operation and the availability of patients for recruitment in time. The sample size obtained for the minimal clinically important difference (MCID) endpoint was feasible and justifiable from an operational and clinical standpoint.

Conclusion:

Evaluation and assessment of multiple endpoints before designing an NI study enable rapid decision-making on endpoint selection and increase operational efficiency.

Keywords: Measuring pain, Pain intensity, Non-inferiority trial, NI margin, Minimal clinically important difference, Sample size estimation.

Article History

Received: August 03, 2022

Revised: January 15, 2023

Accepted: January 16, 2023

1. INTRODUCTION

Disease burden [1] and increasing demand [2] for available drugs continue to drive the search for new treatment options across therapy areas. New innovative and alternative treatment options over standard-of-care treatments add value to overall disease management. The use of non-inferiority trials is an effective strategy adopted [3, 4] by researchers and pharmaceutical companies to assess the benefit of a novel drug compared to a reference drug or standard-of-care. In a non-inferiority study, the intent is to show that the new drug is not inferior to the reference or standard of care by an acceptable margin, which is defined as the non-inferiority margin. The non-inferiority margin is the amount of (inferiority) difference

between a test drug and a reference drug that could have clinically no meaningful difference in efficacy.

In the absence of country-specific regulations and guidance for the industry about the primary endpoint for conducting a clinical trial in a specific therapy area, choosing the right primary endpoint becomes a challenge. Furthermore, in view of risk-benefit comparisons of multiple endpoints, all of which are fairly prevalent and accepted within the clinical community, choosing one primary endpoint is even more difficult. As the primary endpoint has a direct impact on the sample size of the trial and operational costs, it should be carefully and wisely chosen to keep its clinical and scientific merits and risk in mind.

In pain management therapy, the centrally acting, non-opioid analgesic nefopam has demonstrated significantly better

* Address correspondence to this author at the Department of Statistics, Assam University, Silchar, Assam, India; E-mail: anadyabhu@yahoo.com

analgesic activity than placebo in patients with acute postsurgical or fracture pain [5]. In combination with two other non-opioids, paracetamol and ketoprofen, nefopam has demonstrated significant morphine-sparing effect after different types of surgeries, such as cardiac, abdominal, gynecology, orthopedic, and urology, and was found to be associated with superior analgesia in the first 24 h compared to morphine alone [6], indicating an additive analgesic effect. Paracetamol is a non-opiate non-salicylate analgesic with antipyretic and anti-inflammatory properties that blocks the release of certain chemical messengers that cause pain. It is widely and effectively used in both prescription and over-the-counter products to reduce pain and fever [7]. Various studies have reported paracetamol to be effective in controlling pain after oral surgery [6 - 8]. When paracetamol is used in combination with other analgesics, it provides superior pain relief and permits a reduction in opioid dose [9 - 11].

Tramadol is an opioid analgesic that works by blocking the transmission of pain signals to the brain. A fixed-dose combination (FDC) of tramadol and paracetamol [12] is commonly prescribed [13] in post-operative acute pain management and is well-known in family practice for pain relief [13]. Here, we propose a case study involving the design of an inferiority trial to assess the efficacy of an FDC of nefopam + paracetamol versus tramadol + paracetamol in pain management.

Sample size estimation is an important and vital component in planning a clinical trial [14]. For determining the sample size in planning any non-inferiority clinical trial, one of the biggest challenges is to identify a clinically acceptable non-inferiority margin for the considered primary endpoint of the study. Further sample size estimation requires an understanding of the primary objective of the research and its endpoint (measurement scale). This is also critical in the planning of pain management studies, as pain/pain relief is frequently measured and reported on several different endpoints and measurement scales.

1.1. Objective

The objectives of this case study were 1) to explore different sample size estimations for conducting a non-inferiority study on pain management by comparing the efficacy profile of two FDCs nefopam hydrochloride + paracetamol versus tramadol + paracetamol and 2) to choose a suitable endpoint as the primary objective of the study using the proposed sample size options considering study feasibility regarding resourcing and operational costs.

2. MATERIALS AND METHODS

To measure the clinical efficacy in pain management studies, several endpoints have been suggested and frequently reported in the literature, and they are in practice within the medical community [15 - 17]. We were unable to find any specific recommendations on the selection of endpoints for pain measurement from the Indian Health Authority Central Drugs Standard Control Organization (CDSCO). However, the United States Food and Drug Administration, in its draft guidance for industry document, mentioned that “*The primary*

efficacy analysis should compare the sum of pain intensity difference (SPID) between treatments at a prespecified time point that, at a minimum, includes the duration of drug effect, and may extend beyond this duration” [18]. To understand all frequently reported endpoints in pain management studies, an intensive literature search and review were performed. Table 1 provides a summary of the various literature-reported endpoints for the measurement of pain and pain relief. These endpoints are based on the use of continuous scales, such as the visual analog scale (VAS) [19] or a 10-point numerical rating scale (NRS) [20].

As per ICH E10 [21], the margin chosen for a non-inferiority trial cannot be greater than the smallest effect size that the active drug would be reliably expected to have compared with a placebo in the setting of the planned trial. Hence, to understand the non-inferiority margin of identified endpoint, we performed a literature search for preferably placebo-controlled studies with our reference drug (tramadol + paracetamol). We assumed a non-inferiority margin of 20% [22] of the difference. Under proven non-inferiority, this non-inferiority margin will ensure a minimum of 80% of reference treatment effectiveness in our test treatment. Wherever an approximation was needed, we used a conservative approach, which yielded a higher sample size for the given assumption/approximation. The literature reported [22] reference drug and placebo responses, along with their respective non-inferiority margins (derived/proposed), is summarized in Table 2.

2.1. Sample Size Calculation

While calculating the sample size, the overall significance level (α) was maintained at 5%; however, the actual error rate was to be controlled at $\alpha = 0.025$ because of one-sided interest.

For continuous endpoints, the below-mentioned formula was used for the computation of sample size [23]:

$$n = (Z(1-\beta) + Z(1-\alpha))^2 \sigma^2 / [(\mu_A - \mu_B) - dNI]^2$$
 where α is the type I error rate, $(1 - \beta)$ is the power, and Z indicates the critical value of the area under a standard normal distribution.

$(\mu_A - \mu_B)$ is the mean difference between test drug (A) and reference drug (B), and (dNI) is the non-inferiority margin for the respective endpoint.

For dichotomous categorical endpoints, the below-mentioned formula was used for the computation of sample size [24 - 26]:

$$n = (Z(\beta) + Z(\alpha))^2 [PA * (1 - PA) + PB * (1 - PB)] / (PA - PB - dNI)^2,$$

Where α is the type I error rate, $(1 - \beta)$ is the power, and Z indicates the critical value of the area under a standard normal distribution. $(\mu_A - \mu_B)$ is the mean difference between test drug (A) and reference drug (B), and (dNI) is the non-inferiority margin for the respective endpoint.

The estimated sample sizes for each endpoint based on the above-mentioned formulae, literature-reported information about reference drug effectiveness in comparison with placebo, and the calculated NI margins for the respective endpoints, are summarized in Table 3.

2.2. Choice of the Primary Endpoint

The SPID from baseline to 8 hours postdose, SPID(0-8), is the most frequently used endpoint as well as the recommended endpoint by US FDA in its guidance for the industry [18]. Therefore, for our proposed case, we used the SPID (0-8) as the primary endpoint to compare the effectiveness of the FDC.

However, given the available resources and after performing feasibility at hospitals for the availability of subjects willing to participate in a putative clinical trial for the FDC, it was found that the study would not be feasible at a proposed sample size of 1632 with SPID(0-8) as the primary endpoint. Commercial aspects, such as the cost of conducting the study with a large sample size and the market potential of the developed drug to bear the high development cost, also need to be considered when selecting the sample size.

2.3. Alternative Proposal for Primary Endpoints

Table 3 shows that in comparison with continuous endpoints, categorical endpoints require significantly smaller sample sizes while maintaining the same overall error rates and study power. Therefore, even though FDA recommends SPID as the primary endpoint for such a pain management study for FDC development where clinical efficacy is well established

for individual drugs, we proposed meaningful pain relief (MPAR) with minimum clinically important difference MCID) [27 - 29] as the primary endpoint, and SPID as the key secondary endpoint.

3. RESULTS

The sample size requirement depends on the nature of sensitivity and variation of the chosen primary objective and corresponding endpoint for the study. The measurement of pain became complex and composite when it was measured with respect to time points. The different endpoints used to measure pain, pain intensity, and pain improvement are summarized in Table 1.

For the above-mentioned different endpoints, the non-inferiority margins are derived considering literature that reported a placebo treatment effect and a 20% of acceptable inferiority margin. The derived non-inferiority margins for each endpoint are summarized in Table 2.

With the help of literature reported treatment effect (the control treatment/reference treatment) and expected treatment effect of our test treatment and with the above derived non-inferiority margins of each endpoint, the sample size was derived for each endpoint (Table 3).

Table 1. Endpoints to measure pain and pain relief.

Endpoint	Measure	Definition
Continuous scales: VAS**** or 10-point NRS**		
Pain intensity at time t	PI (t)	Pain intensity measured at time t. Directedly responded by subject experiencing pain and recorded at the source.
Pain intensity difference from baseline at time t	PID (t)	Derived by subtracting pain intensity at each time point t from pain intensity at baseline time 0: $-[PI(t) - PI(0)]$.
Sum of pain intensity difference at time t	SPID (t)***	Derived by multiplying the PID score at each post-dose time point (Ti) by the duration (in hours) since the preceding time point (Ti-1) and then summing the values over the observation period: $SPID-t = \sum_{i=1}^t Ti - Ti-1 * PIDi$
Percentage of maximum sum of pain intensity difference at time t	% of Max SPID (t)	Derived from the ratio of max of SPID by SPID at time t multiplied by 100: $Max\ of\ [Ti - Ti-1 * PIDi \ \square \ i \leq t] \ [SPID-t] * 100$
Pain relief at time t	PAR (t)	Pain relief measured at time t. Directedly responded by subject experiencing pain and recorded at the source.
Maximum pain relief	Max PAR	Derived as maximum pain relief experienced during the observation: $Max\ of\ (PARi \ \square \ i \leq t)$
Total pain relief at time t	TOTPAR (t)	Derived by multiplying the PAR score at each post-dose time point by the duration (in hours) since the preceding time point and then summing the values over the observation period: $TOTPAR-t = \sum_{i=1}^t Ti - Ti-1 * PARi$
Percentage of maximum total pain relief at time t	% of max TOTPAR (t)	Derived as the ratio of max of SPID by SPID at time t multiplied by 100 $Max\ of\ [Ti - Ti-1 * PARi \ \square \ i \leq t] \ [TOTPAR-t] * 100$
Categorical scale: Mostly dichotomous in % of responders		
At least 50% of max TOTPAR	50% of max TOTPAR (t)	Derived as a percentage of subjects experiencing at least 50% of maximum total pain relief
At least 30% PI reduction	30% of max TOTPAR (t)	Derived as a percentage of subjects experiencing at least 30% of maximum total pain relief
Meaningful pain relief/MCID	MPAR (t)	Derived as a percentage of subjects experiencing meaningful pain reduction based on MCID* criteria

Abbreviations: *MCID, minimal clinically important difference; **NRS, numerical rating scale; ***SPID, sum of pain intensity difference; ****VAS, visual analog scale

Table 2. Derived non-inferiority margins for each pain management endpoint.

Endpoint / Assessment Parameter	Scale	Tramadol + Paracetamol	Placebo	Difference (Maximum of Treatment Effect) Margin M1	20% of the Difference as an Acceptable NI Margin
Continuous scale: Mostly on Visual Analog Scale (VAS) or Numeric Rating Scale (NRS): Mean (SD)					
PI (8)	VAS	5.5 (1.3)	5.2 (1.3)	0.3	0.06
PID (8)	VAS	1.0 (1.5)	0.3 (1.5)	0.7	0.14
SPID (8)	VAS	14.4 (15.21)	2.2 (8.98)	12.2	2.44
% of Max SPID (8)	VAS	34.1 (34.89)	5.2 (21.48)	28.9	5.78
PAR (t)	VAS	2.2 (1.5)	0.5 (1.2)	1.7	0.34
Max PAR	VAS	2.9 (1.5)	2 (1.2)	0.9	0.18
TOTPAR (t)	NRS	9.2 (7.65)	1.9 (3.89)	7.3	1.46
% of Max TOTPAR (8)	NRS	34.2 (30.76)	7.1 (15.86)	27.1	5.42
Categorical scale: Mostly dichotomous in % of responders					
50% of Max TOTPAR (8)	NRS	35.50%	5.30%	30.2	20
30% of Max TOTPAR (8)	NRS	35.50%	9.90%	25.6	20
MPAR (8)	VAS	72.90%	25.20%	47.7	20

Abbreviations: NI, Non-Inferiority; SD, Standard Deviation

Table 3. Recommended sample sizes for each pain management endpoint.

Endpoint	Estimated Sample Size
PI (8)	19734
PID (8)	4828
SPID (8)	1632
% of Max SPID (8)	1534
PAR (t)	820
Max PAR	2922
TOTPAR (t)	1156
% of Max TOTPAR (8)	1356
50% of Max TOTPAR (8)	242
30% of Max TOTPAR (8)	242
MPAR (8)	728

With the alternative proposal, which demands a sample size of a minimum of 728 evaluable subjects with 90% power, which is less than half of the sample size of 1632 needed with SPID as the primary endpoint, a visual comparison of the recommended sample size against each endpoint is made on a semi-log scale, as shown in Fig. (1).

The alternative proposed sample size was reasonably acceptable for the proposed case (both drugs in their independent form have well-proven efficacy, are available in the market, and the fixed-dose combination of both needs to be non-inferior to their free form) considering resourcing and operational costs involved in the development of FDC vs. the actual benefit and added convenience to the patients.

However, the risk and benefits of such a proposal should be considered in terms of its acceptance by the clinical community and regional regulatory authorities and benefits to subjects participating in the clinical study.

4. DISCUSSION

In view of the proposed case, especially with a non-

inferiority study design, the development of an FDC had the primary aim of adding efficacy to treatment management. When the effectiveness and safety of individual drug components are well proven, the research and development costs of a non-inferior FDC with a very high sample size have the potential to increase the cost of a drug available to patients, thus questioning the rationale of such FDC development.

As per the general statistical principle for fixed sample size, making any inference on continuous data point is more powerful than discrete, which means that for our primary endpoint, we may need a smaller sample size on continuous outcome over a discrete outcome. However, the FDA-recommended endpoint SPID is a derived composite and needs time, time intervals, and multiple assessments of pain intensity difference. This makes SPID, as the primary endpoint, more sensitive and accurate, containing the entire spectrum of pain profiles with respect to time. Simultaneously, due to its statistical sensitivity and variability in nature, the sample size estimate on this endpoint tends to be larger. It may become non-feasible for trial overall clinical operations and cost-effective for non-inferiority study purposes.

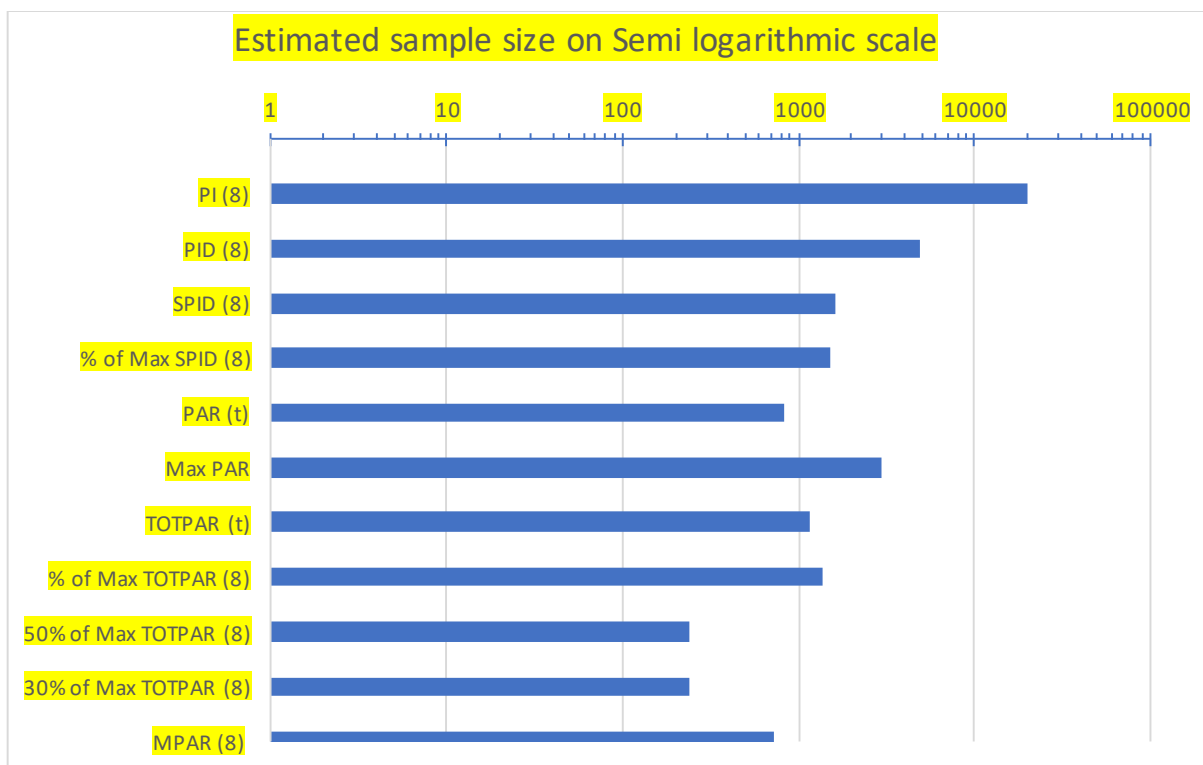


Fig. (1). Recommended sample sizes for each pain management endpoint on a semi-logarithmic scale.

The alternate proposal of choosing the proportion of subjects achieving MPAT (t) with MCID criteria as the primary endpoint required a significantly lower (lesser than half) sample size (728) in comparison to the first proposal. With this sample size for the primary endpoint and SPID as the key secondary endpoint, it is expected that optimal information may be obtained about the entire spectrum of pain profiles during treatment.

The endpoint MPAT (t), with MCID criteria as the primary endpoint, demands a very low sample size compared to SPID and serves the statistical purpose of achieving good power with a lesser sample size. It has its own limitations. It only assesses at the end of the study how many subjects received clinically meaningful pain relief. Concerning pain relief at an interim time point, this cannot be a very suitable endpoint. The use of such an endpoint should always be in alignment with the clinical context of relevance. Moreover, it should be accompanied by other continuous endpoints like PID and SPID as key secondary objectives for a better understanding of the treatment profile.

The limitation of the present study is that the information needed to estimate the sample size was not available from previous/pilot studies for the test treatment. We selected those studies where all mentioned endpoints were reported to be consistent and comparable in the sample size proposal.

CONCLUSION

A clinically acceptable endpoint and a well-powered adequate sample size for the selected endpoint are important

and essential criteria in the design of clinical trials. The dependency of sample size on the selected endpoint and the chosen non-inferiority margin increase the complexity of sample size estimation.

In conclusion, the study aimed to propose a clinically acceptable endpoint for evaluating the effectiveness of an FDC of nefopam + paracetamol versus tramadol + paracetamol in managing pain. SPID (0-8), feasibility of SPID (0-8), and MPAR with MCID criteria were determined, and an endpoint with a reasonably feasible sample size considering operational, resource, and development costs was chosen. Such a strategy may be employed in determining endpoints and sample sizes for similar studies in pain management.

LIST OF ABBREVIATIONS

- FDC** = fixed-dose combination
- CDSCO** = central drugs standard control organization
- SPID** = sum of pain intensity difference
- VAS** = visual analog scale

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No animals/humans were used in the studies that are the basis of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The data supporting the findings of the article is available in public domain at mentioned below URL.

1. <https://bmjopen.bmj.com/content/9/2/e023715>
2. <https://bmcanaesthesiol.biomedcentral.com/articles/10.1186/s12871-016-0174-5>.

FUNDING

None.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

- [1] Truglio J, Graziano M, Vedanthan R, *et al.* Global health and primary care: Increasing burden of chronic diseases and need for integrated training. *Mt Sinai J Med* 2012; 79(4): 464-74. [<http://dx.doi.org/10.1002/msj.21327>] [PMID: 22786735]
- [2] Giorgia S, Batomen B, Kotwani A, Pai M, Gandra S. Sales of antibiotics and hydroxychloroquine in India during the COVID-19 epidemic: An interrupted time series analysis. *PLoS Med* 18(7): e1003682.2021; [<http://dx.doi.org/10.1371/journal.pmed.1003682>]
- [3] Head SJ, Kaul S, Bogers AJJC, Kappetein AP. Non-inferiority study design: Lessons to be learned from cardiovascular trials. *Eur Heart J* 2012; 33(11): 1318-24. [<http://dx.doi.org/10.1093/eurheartj/ehs099>] [PMID: 22564354]
- [4] FDA guidance for industry documents. Non-inferiority clinical trials to establish effectiveness guidance for industry. 2016. <https://www.fda.gov/media/78504/download>
- [5] Wang RH, Waite EM. The clinical analgesic efficacy of oral nefopam hydrochloride. *J Clin Pharmacol* 1979; 19(7): 395-402. [<http://dx.doi.org/10.1002/j.1552-4604.1979.tb02498.x>] [PMID: 479385]
- [6] Weil K, Hooper L, Afzal Z, *et al.* Paracetamol for pain relief after surgical removal of lower wisdom teeth. *Cochrane Libr* 2007; 2007(3): CD004487. [<http://dx.doi.org/10.1002/14651858.CD004487.pub2>] [PMID: 17636762]
- [7] Anderson BJ. Paracetamol (Acetaminophen): Mechanisms of action. *Paediatr Anaesth* 2008; 18(10): 915-21. [<http://dx.doi.org/10.1111/j.1460-9592.2008.02764.x>] [PMID: 18811827]
- [8] Kiersch TA, Halladay SC, Hormel PC. A single-dose, double-blind comparison of naproxen sodium, acetaminophen, and placebo in postoperative dental pain. *Clin Ther* 1994; 16(3): 394-404. [PMID: 7923306]
- [9] Seymour RA, Hawkesford JE, Sykes J, Stillings M, Hill CM. An investigation into the comparative efficacy of soluble aspirin and solid paracetamol in postoperative pain after third molar surgery. *Br Dent J* 2003; 194(3): 153-7. [<http://dx.doi.org/10.1038/sj.bdj.4809898>] [PMID: 12598885]
- [10] Toms L, Derry S, Moore RA, McQuay HJ. Single dose oral paracetamol (acetaminophen) with codeine for postoperative pain in adults. *Cochrane Libr* 2009; 2019(5): CD001547. [<http://dx.doi.org/10.1002/14651858.CD001547.pub2>] [PMID: 19160199]
- [11] Miranda HF, Puig MM, Prieto JC, Pinardi G. Synergism between paracetamol and nonsteroidal anti-inflammatory drugs in experimental acute pain. *Pain* 2006; 121(1): 22-8. [<http://dx.doi.org/10.1016/j.pain.2005.11.012>] [PMID: 16480830]
- [12] Ciurba A, Hancu G, Cojoclea LM, Sipos E, Todoran N. Development of new formulation and its evaluation by capillary electrophoresis of tablets containing tramadol hydrochloride and paracetamol. *Pharm Dev Technol* 2014; 19(7): 833-8. [<http://dx.doi.org/10.3109/10837450.2013.836219>] [PMID: 24028179]
- [13] Tramadol/paracetamol fixed-dose combination for chronic pain management in family practice: A clinical review. *Int Sci Res Not* 2013. Article ID: 638469 [<http://dx.doi.org/10.5402/2013/638469>]
- [14] Biau DJ, Kernéis S, Porcher R. Statistics in brief: The importance of sample size in the planning and interpretation of medical research. *Clin Orthop Relat Res* 2008; 466(9): 2282-8. [<http://dx.doi.org/10.1007/s11999-008-0346-9>] [PMID: 18566874]
- [15] Englbrecht M, Tarner I, Manger B, Bombardier C, Müller-ladner U, Heijde V. Measuring pain and efficacy of pain treatment in inflammatory arthritis: A systematic literature review. *The Journal of Rheumatology Supplement* 2012; 90: 3-10. [<http://dx.doi.org/10.3899/jrheum.120335>]
- [16] Miles CL, Pincus T, Carnes D, Taylor SJ, Underwood M. Measuring pain self-efficacy. *Clin J Pain* 2011; 27(5): 461-70. [<http://dx.doi.org/10.1097/AJP.0b013e318208c8a2>] [PMID: 21317778]
- [17] Amtmann Dagmar, Liljenquist Kendra, Bamer Alyssa, *et al.* Measuring pain catastrophizing and pain-related self-efficacy: Expert panels, focus groups, and cognitive interviews *Patient* 2018; 11(i): 107-17.
- [18] U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER). Development of non opioid analgesics for acute pain : Guidance for industry draft guidance. 2022. <https://www.fda.gov/media/156063/download>
- [19] Hawker GA, Mian S, Kendzerska T, French M. Measuring of adult pain. *Arthritis Care Res* 2011; 63: S240-52. [<http://dx.doi.org/10.1002/acr.20543>] [PMID: 22588748]
- [20] Iohom G. Post Operative Pain Management An Evidence based Guide to Practice, Clinical Assessment of Post operative Pain. *ScienceDirect. ScienceDirect* 2006; Chapter 11: pp. 102-8. [<http://dx.doi.org/10.1016/B978-1-4160-2454-5.50016-3>]
- [21] Choice of control group and related Issues in clinical trials. International conference on harmonisation of technical Requirements for registration of pharmaceuticals for human Use (ICH) harmonized tripartite guideline (2000) E10. https://database.ich.org/sites/default/files/E10_Guideline.pdf
- [22] Gay-Escoda C, Hanna M, Montero A, *et al.* Tramadol/dextketoprofen (TRAM/DKP) compared with tramadol/paracetamol in moderate to severe acute pain: results of a randomised, double-blind, placebo and active-controlled, parallel group trial in the impacted third molar extraction pain model (DAVID study). *BMJ Open* 2019; 9(2): e023715. [<http://dx.doi.org/10.1136/bmjopen-2018-023715>] [PMID: 30782886]
- [23] Beal SL. Sample size determination for confidence intervals on the population mean and on the difference between two population means. *Biometrics* 1989; 45(3): 969-77. [<http://dx.doi.org/10.2307/2531696>] [PMID: 2790131]
- [24] Fleiss JL, Tytun A, Ury HK. A simple approximation for calculating sample sizes for comparing independent proportions. *Biometrics* 1980; 36(2): 343-6. [<http://dx.doi.org/10.2307/2529990>] [PMID: 26625475]
- [25] Chow S-C, Shao J, Wang H. *Sample Size Calculations in Clinical Research*. Boca Raton, FL: CRC Press 2003; p. 358. [<http://dx.doi.org/10.1201/9780203911341>]
- [26] Sample size calculator: Two parallel-sample proportions. https://www2.ccrb.cuhk.edu.hk/stat/proportion/tssp_sup.htm
- [27] Jaeschke R, Singer J, Guyatt GH. Measurement of health status. *Control Clin Trials* 1989; 10(4): 407-15. [[http://dx.doi.org/10.1016/0197-2456\(89\)90005-6](http://dx.doi.org/10.1016/0197-2456(89)90005-6)] [PMID: 2691207]
- [28] Myles PS, Myles DB, Gallagher W, *et al.* Measuring acute postoperative pain using the visual analog scale: The minimal clinically important difference and patient acceptable symptom state. *Br J Anaesth* 2017; 118(3): 424-9. [<http://dx.doi.org/10.1093/bja/aew466>] [PMID: 28186223]
- [29] Olsen MF, Bjerre E, Hansen MD, *et al.* Pain relief that matters to patients: systematic review of empirical studies assessing the minimum clinically important difference in acute pain. *BMC Med* 2017; 15(1): 35. [<http://dx.doi.org/10.1186/s12916-016-0775-3>] [PMID: 28215182]