



Thèse

2020

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Utilisation des échelles de douleur dans les études péri-opératoires chez l'enfant : Une analyse transversale et une enquête auprès des auteurs

Gerstman, Michelle Diana

How to cite

GERSTMAN, Michelle Diana. Utilisation des échelles de douleur dans les études péri-opératoires chez l'enfant : Une analyse transversale et une enquête auprès des auteurs. Doctoral Thesis, 2020. doi: 10.13097/archive-ouverte/unige:150925

This publication URL: <https://archive-ouverte.unige.ch/unige:150925>

Publication DOI: [10.13097/archive-ouverte/unige:150925](https://doi.org/10.13097/archive-ouverte/unige:150925)

Section de Médecine Clinique
Département APSIU
(Anesthésiologie, Pharmacologie, Soins
Intensifs et Urgences)
Service d'Anesthésiologie

Thèse préparée sous la direction du Professeur Martin Tramèr et de la Dre Nadia Elia, PD

Utilisation des échelles de douleur dans les études péri-opératoires chez l'enfant

Une analyse transversale et une enquête auprès des auteurs

Thèse

présentée à la Faculté de Médecine

de l'Université de Genève

pour obtenir le grade de Docteur en médecine

par

Michelle Diana GERSTMAN

de

Adelaide (Australie)

Thèse n° 10974

Genève

2020

Table de matières

1	<i>Remerciements</i>	4
2	<i>Présentation</i>	5
3	<i>Déclaration de conflit d'intérêt et financements</i>	6
4	<i>Glossaire des abréviations</i>	7
5	<i>Résumé</i>	8
6	<i>Introduction</i>	9
6.1	La douleur	9
6.1.1	Définition.....	9
6.1.2	La douleur chez les enfants.....	9
6.1.3	Les échelles de douleur	10
6.1.4	Les caractéristiques d'une échelle de douleur idéale chez l'enfant.....	11
6.1.5	Types des échelles de douleur chez les enfants.....	11
6.2	La recherche sur la prévention et le traitement de la douleur chez l'enfant ...	13
6.2.1	La revue systématique	13
6.2.2	PedIMMPACT	14
6.2.3	Critères de qualité des revues systématiques et des scores selon PedIMMPACT	16
6.2.4	Domaines principaux de résultats	17
6.2.5	Les recommandations de mesures de la douleur	18
6.2.6	PedIMMPACT et l'utilisation des échelles de la douleur dans la recherche	21
7	<i>Objectifs de ce travail</i>	22
8	<i>Methods</i>	23
8.1	Study design	23
8.2	Setting	23
8.2.1	Trials and search strategy	23
8.2.2	Inclusion criteria.....	23
8.2.3	Exclusion criteria	24
8.3	Variables	24
8.4	Data sources/measurement	24

8.5	Bias	25
8.6	Study size	25
8.7	Quantitative variables	25
8.8	Statistical methods	26
9	Results	28
9.1	Trial selection	28
9.2	Descriptive data	28
9.2.1	Pain scores used	29
9.3	Outcome data.....	30
9.3.1	Primary outcomes	30
9.3.2	Secondary outcomes	32
10	Discussion	33
10.1	Key results.....	33
10.2	Interpretation.....	33
10.3	Strengths and limitations.....	35
10.4	Generalizability	37
11	Bibliography.....	38
12	Figures	41
13	Tableaux	48
14	Annexe: Références bibliographiques des articles inclus dans l'analyse transversale.....	59

1 Remerciements

Je voudrais remercier mon directeur de thèse le Prof Martin Tramèr pour son soutien tout au long de mon travail.

Je voudrais également exprimer toute ma gratitude à Dre Nadia Elia. Toujours à l'écoute, j'ai énormément appris à ses côtés tant au niveau des conseils méthodologiques que des analyses statistiques et la démarche scientifique.

Je voudrais aussi remercier les co-auteurs de l'article qui est à la base de cette thèse, le Prof Walid Habre et la Dre Lucie Rolland, qui ont contribué à la réalisation de ce travail scientifique.

Je voudrais également remercier ma famille et mes amis pour leur soutien pendant ce travail.

2 Présentation

J'ai débuté ce travail avec l'idée de faire une revue systématique avec méta-analyse concernant l'intérêt que présentait l'utilisation d'un opiacé, le tramadol, comme antalgique en période péri-opératoire chez les enfants. Après avoir commencé le travail, en analysant les études publiées, il est devenu assez vite clair qu'une méta-analyse ne serait pas possible à cause de toutes les différentes échelles de douleurs utilisées par les chercheurs. A ce moment, ma question de recherche s'est transformée et je me suis intéressée à comprendre comment les chercheurs choisissaient une échelle de douleur plutôt qu'une autre pour leurs études expérimentales chez l'enfant. C'est donc sur ce sujet que ma thèse va porter.

Cette thèse est présentée avec, en premier, un résumé, une introduction et les objectifs qui sont rédigés en français. Par la suite, je présente en anglais les chapitres *Methods*, *Results* et *Discussion* qui font l'objet d'un article qui est actuellement en cours de soumission dans un journal scientifique avec revue par un comité de lecture sous le titre suivant :

Gerstman M, Rolland L, Tramèr M, Habre W, Elia N. *The use of pain scores in pediatric research studies. A cross-sectional analysis and survey of authors.*

Suivent enfin la bibliographie, les tableaux et figures ainsi que, comme annexe, toutes les références bibliographiques des articles inclus dans l'analyse transversale (en ordre alphabétique).

3 Déclaration de conflit d'intérêt et financements

Je n'ai pas des conflits d'intérêts à déclarer en lien avec cette thèse. En particulier, ce travail n'a pas bénéficié de financement externe.

4 Glossaire des abréviations

AIIMS	<i>All India Institute of Medical Sciences</i>
CHEOPS	<i>Children's Hospital of Eastern Ontario Pain Scale</i>
CHIPPS	<i>Children's and Infants' Postoperative Pain Scale</i>
ENT	<i>Ear Nose and Throat</i>
FDA	<i>US Food and Drug Administration</i>
FLACC	<i>Face, Legs, Activity, Crying, Consolability</i>
FPSR	<i>Faces Pain Scale Revised</i>
IASP	<i>International Association for the Study of Pain</i>
NRS	<i>Numerical Rating Scale</i>
PedIMMPACT	<i>Pediatric Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials</i>
PPPM	<i>Parents' Postoperative Pain Measure</i>
RCT	<i>Randomized Controlled Trial</i>
TPPPS	<i>Toddler-Preschooler Postoperative Pain Scale</i>
VAS	<i>Visual Analog Scale</i>

5 Résumé

Objectif : Premièrement, comprendre les raisons qui sous-tendent le choix d'une échelle d'évaluation de la douleur particulière par les chercheurs dans les essais cliniques portant sur la douleur péri-opératoire en pédiatrie. Deuxièmement, vérifier si les recommandations internationales sur l'évaluation de la douleur, PedIMMPACT (Pediatric Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials), issue d'un consensus d'experts, ont impacté le choix des échelles de douleur dans ces études.

Design : Analyse transversale avec recueil systématique du choix des échelles de douleur par les auteurs de tous les essais contrôlés randomisés en péri-opératoire publiés entre 2000 et 2017 portant sur l'administration du tramadol chez les enfants de 1 mois à 16 ans, et enquête par questionnaire auprès des auteurs lorsque leur choix n'était pas décrit dans l'essai publié.

Résultats : Sur les 59 études identifiées et analysées, 13 expliquaient leur choix dans l'article publié et incluaient une référence, 31 fournissaient une référence mais sans donner de justification pour leur choix et 15 ne fournissaient ni description ni référence. Le choix d'échelle de la douleur utilisée par les chercheurs était le plus souvent motivé par le fait qu'elle était « validée » (31%), ou « validée et familière » (23%). Les références fournies étaient majoritairement des « études préalables » utilisant la même échelle (40%), ou des études « validant » une échelle de la douleur (38%). Un seul essai a justifié son choix par les recommandations PedIMMPACT. Il n'y a pas eu d'augmentation du nombre d'essais utilisant des échelles recommandées par PedIMMPACT après sa publication.

Conclusion : Le choix des échelles de douleur utilisées durant la période péri-opératoire dans les études pédiatriques semble aléatoire et une grande variété d'échelles sont utilisées. Les recommandations PedIMMPACT ne sont pas appliquées.

6 Introduction

6.1 La douleur

6.1.1 Définition

La douleur est définie par l'association Internationale d'Étude de la Douleur (*International Association for the Study of Pain – IASP*) comme « une sensation et une expérience émotionnelle désagréable en réponse à une atteinte tissulaire réelle ou potentielle ou décrites en ces termes » (1).

Elle résulte d'un phénomène complexe qui commence par une stimulation nociceptive via les nerfs périphériques et puis un processus central qui intègre ces informations différemment selon l'expérience, la personnalité et les prédispositions génétiques de la personne (2). Le sentiment et la sensation de douleur varient donc entre chaque personne. Il en est de même de l'effet des médicaments antalgiques.

La douleur survient à différentes occasions au cours de la vie et endosse généralement un rôle protecteur. Parfois pourtant, la douleur peut se chroniciser et devenir une pathologie en elle-même. Certaines organisations proposent que la douleur soit considérée comme un « signe vital » (3). Qu'elle soit aiguë ou chronique, les prises en charge et les traitements visant à soulager la douleur sont essentiels, et leur accès est même considéré comme un des droits de l'homme (4).

6.1.2 La douleur chez les enfants

Dans les années 70, plusieurs auteurs ont commencé à s'inquiéter du fait que la douleur chez l'enfant était peu étudiée, et souvent n'était pas traitée de manière optimale (5, 6).

Ceci était d'autant plus dommageable qu'il a été démontré que l'expérience de la douleur chez l'enfant pouvait avoir des effets néfastes à court, mais aussi à long terme. En effet, des expériences douloureuses durant l'enfance peuvent rendre la personne plus sensible à la douleur plus tard, et peuvent également avoir un effet

« traumatique » (7). Puisqu'une intervention chirurgicale est un moment où un enfant est à risque d'être exposé à la douleur, c'est une période durant laquelle il est important de tout mettre en œuvre afin de pouvoir la contrôler et minimiser ses conséquences potentielles à court comme à long terme.

L'absence d'études sur la douleur chez l'enfant a été mise sur le compte de la difficulté de *mesurer* la douleur dans cette population particulière et, malgré certains progrès effectués au cours de ces 50 dernières années, ce domaine de l'évaluation et la prise en charge de la douleur chez l'enfant continue d'être un champ nécessitant des améliorations (8-10).

6.1.3 Les échelles de douleur

Les échelles, parfois également appelées « scores », de douleur sont des outils permettant de mettre en évidence la présence d'une douleur, de quantifier son intensité, mais également permettant de mesurer l'efficacité d'une mesure de prévention ou l'effet antalgique d'un traitement (11). L'utilisation de ces scores, en clinique, vise à améliorer la prise en charge de la douleur et la satisfaction des patients (12) alors que dans le domaine de la recherche, les scores de douleur utilisés doivent permettre de mettre en évidence des différences parfois petites dans l'efficacité antalgique de différents médicaments ou prises en charge. Il est donc très important que les scores utilisés pour la recherche en antalgie soient fiables, comparables et reproductibles *entre* les enfants. Dans le contexte clinique par contre, il est surtout important que le score soit facile d'utilisation (par l'enfant et par les soignants) et reproductible dans le temps, chez le même enfant.

Néanmoins, dans un monde idéal, les mêmes mesures de douleurs devraient être utilisées dans la recherche que dans la clinique. Cela contribuerait à rendre les résultats des essais cliniques plus applicables à la pratique clinique. (13)

Un problème particulier chez les enfants, est qu'il est particulièrement difficile de mesurer leur douleur (14). En effet, les très jeunes enfants n'ont pas encore la capacité de s'exprimer clairement, et même quand ils grandissent, leur capacité d'intégrer les entrées nociceptives, de les comprendre, puis de les exprimer n'est pas encore très bien développée.

C'est une des raisons pour laquelle de multiples échelles de la douleur se sont développées en parallèle, dont le but était de tenter de quantifier au mieux l'intensité de la douleur chez l'enfant.

6.1.4 Les caractéristiques d'une échelle de douleur idéale chez l'enfant

Certains auteurs ont tenté de définir les caractéristiques nécessaires à un outil de mesure de la douleur chez l'enfant (15). Ils ont retenu 10 points.

Une échelle idéale devrait :

1. Être valide et fiable. Sa validité et sa fiabilité devraient être prouvées scientifiquement. Ces concepts de validité et de fiabilité seront développés plus loin.
2. Convenir au stade de développement et à la langue maternelle de l'enfant.
3. Être facile à comprendre par les enfants (ou leur parents), même par ceux qui ont un niveau d'éducation bas.
4. Être facile et rapide à expliquer par les soignants aux enfants ou à leurs parents.
5. Être facile et rapide à utiliser, par les soignants.
6. Être appréciée par les parents, les enfants et les soignants.
7. Être utilisable, avec les enfants, pour fixer des objectifs de traitement.
8. Être pas chère et facile à obtenir, à reproduire et à distribuer.
9. Être facilement désinfectée.
10. Être disponible dans plusieurs langues.

6.1.5 Types des échelles de douleur chez les enfants

Il existe 2 groupes principaux d'échelles de la douleur qui sont utilisées chez les enfants: les échelles d'auto-évaluation et les échelles comportementales (15). Un troisième groupe, celui des échelles physiologiques existe également, mais ces échelles sont considérées comme trop « indirectes » et si elle peuvent présenter un certain intérêt en accompagnement des autres échelles, elles ne sont en général pas utilisées comme échelles principales de mesure de la douleur en clinique (15).

En règle générale, la préférence va à l'utilisation des échelles d'autoévaluation. Cependant, quand elles se montrent inadaptées au contexte ou à l'enfant, ou que leur utilisation est impossible, les échelles comportementales sont utilisées.

6.1.5.1 Les échelles d'auto-évaluation

Les échelles d'auto-évaluation visent à permettre à l'enfant d'exprimer lui-même l'intensité de la douleur ressentie (16). La douleur étant subjective, les échelles d'autoévaluation sont généralement considérées comme le meilleur moyen d'évaluation de la douleur (en anglais le *gold standard*) (17). Elles se présentent sous différentes formes, par exemple sous la forme de chiffres de 0 (représentant pas de douleur du tout) à 10 (la pire des douleurs imaginables), ou à l'aide d'une suite de mots comme, par exemple de « pas mal du tout » à « très mal », ou encore avec une succession de photos, ou de dessins de visages montrant différentes émotions liées à l'expérience de la douleur (figure 1) (18).

L'âge à partir duquel un enfant est capable de comprendre et d'utiliser une échelle d'autoévaluation varie considérablement selon les enfants et selon les experts. On considère, en général qu'à partir de 5 ans, la plupart des enfants parviennent à les utiliser. Cette capacité ne varie pas seulement avec l'âge, mais aussi selon le contexte. Par exemple, lors d'un épisode de maladie, le stade de développement d'un enfant peut « régresser ». Il a également été suggéré que la capacité d'un enfant à utiliser un score d'auto-évaluation augmente avec l'entraînement, c'est-à-dire s'il a déjà été « exposé » précédemment à l'utilisation de ce score (15). Enfin, la fréquentation scolaire semble avoir également une incidence sur la capacité d'un enfant à utiliser les mesures d'autoévaluation (19). Une variété d'outils a été proposée pour tenter de déterminer quels enfants sont capables d'utiliser une mesure d'auto-évaluation et lesquels ne le sont pas. On utilise par exemple des vignettes, des exercices de classement d'objets ou de formes, on évalue la compréhension de l'enfant, et on mène des discussions autour de l'expérience antérieure de l'enfant en matière de douleur (20). Un exemple d'exercice de classement est la « sériation » : on donne six triangles de tailles différentes à un enfant et on lui demande de sélectionner le plus grand triangle et le plus petit triangle, que l'on met de côté. On demande alors à l'enfant de sélectionner, à

nouveau, le plus grand des triangles restant, et ainsi de suite jusqu'au moment où il ne reste plus de triangles. Si l'enfant est capable de faire cet exercice, de classer les objets, ou de parler de son expérience de la douleur, il est présumé que cet enfant pourra utiliser une échelle d'auto-évaluation. Malgré tout, la capacité de discrimination de ces outils reste mal établie et l'âge, bien que non parfait, semble être le meilleur prédicteur de la capacité d'un enfant à utiliser une échelle d'auto-évaluation de la douleur. (21)

6.1.5.2 Les échelles comportementales

Les échelles comportementales sont les échelles qui permettent à quelqu'un d'autre que l'enfant lui-même d'estimer la douleur ressentie par un enfant, souvent lorsque ce dernier n'est pas en mesure d'utiliser une échelle d'auto-évaluation. Typiquement, il s'agit d'un soignant ou plus rarement, un parent (16). Les éléments d'évaluation utilisés dans les échelles comportementales comprennent, par exemple, la présence ou l'absence de pleurs, d'un état d'agitation, la capacité de l'enfant à être consolé, la présence de mouvements, la posture de l'enfant, ou encore la localisation de la douleur et l'expression faciale de l'enfant.

Les échelles comportementales sont utilisées chez les très jeunes enfants (généralement moins que 4 ans), ainsi que chez les enfants incapables de s'exprimer par auto-évaluation pour diverses raisons, par exemple les enfants limités par des problèmes cognitifs, physiques (des bandages qui les empêchent de parler) ou limités dans leur capacité par les traitements reçus (19).

6.2 La recherche sur la prévention et le traitement de la douleur chez l'enfant

6.2.1 La revue systématique

La recherche clinique est nécessaire à l'amélioration des soins aux patients en général, et dans le domaine de la douleur également. Les études expérimentales,

randomisées et contrôlées, représentent les études de référence pour comparer les différentes méthodes de prévention et/ou de traitement de la douleur et ainsi permettre de choisir en conséquence les soins les plus appropriés. La revue systématique accompagnée d'une analyse quantitative appelée « méta-analyse » combine les résultats de différentes études. Elle est considérée comme le design d'étude fournissant le meilleur niveau de preuve de l'efficacité et/ou de la supériorité d'une prise en charge ou d'un traitement spécifique sur les traitements habituels (22).

Cependant, pour être probante, une revue systématique accompagnée d'une méta-analyse doit se baser sur différentes études présentant un minimum d'homogénéité. L'obtention d'un résultat quantitatif fiable est parfois rendu difficile voire impossible par l'utilisation d'échelles de mesure de la douleur très différentes d'une étude à l'autre. Par exemple, il est difficile de comparer une intensité de la douleur de 7 sur 13 évaluée par une échelle type CHEOPS (*Children's Hospital of Eastern Ontario Pain Scale*), avec une de 4 sur 10 obtenue avec l'échelle FLACC (*Face, Legs, Activity, Crying, Consolability*).

Étant donné la grande variété des échelles de la douleur qui existent chez l'enfant, il est particulièrement difficile de tirer des conclusions concernant les meilleures procédures de prévention ou de traitement de la douleur en se basant sur le meilleur niveau de preuve. C'est dans ce contexte qu'en 2008, une déclaration de consensus (en anglais : *consensus statement*) a vu le jour qui avait pour but d'uniformiser l'utilisation des échelles de la douleur dans les études interventionnelles, chez l'enfant. Cette déclaration a été publiée par le groupe connu aujourd'hui sous le nom de PedIMMPACT.

6.2.2 PedIMMPACT

The Pediatric Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (PedIMMPACT) avait comme double objectif d'augmenter le nombre d'études sur la prise en charge de la douleur chez l'enfant d'une part, et d'améliorer la capacité à comparer ces études entre elles d'autre part (23). Pour y répondre,

PedIMMPACT a développé et décrit les « domaines principaux de résultats et de mesures » utiles à rapporter dans les études de douleur chez l'enfant.

Ce groupe d'experts composé de 26 professionnels issus de milieux universitaires, gouvernementaux et de l'industrie pharmaceutique, s'est réuni pour la première fois en 2005.

Le groupe PedIMMPACT a d'abord mandaté l'élaboration de 2 revues systématiques portant sur la prise en charge de la douleur chez l'enfant. La première revue systématique s'est concentrée sur les échelles d'auto-évaluation afin de définir lesquelles devaient être recommandées (24), alors que la deuxième revue systématique s'est focalisée sur les échelles comportementales (19). Ces 2 revues systématiques ont été publiés en 2006.

Par la suite, le groupe s'est intéressé au développement des « domaines principaux de résultats » en lien avec la douleur. Après discussions au sein de ce groupe et consultations avec la communauté scientifique s'intéressant à la prise en charge de la douleur chez l'enfant, PedIMMPACT a défini 6 « domaines principaux de résultats ». Ils les ont publiés en 2008 dans une déclaration de consensus : *Consensus statement «Core Outcome Domains and Measures for Pediatric Acute and Chronic/Recurrent Pain Clinical Trials: PedIMMPACT Recommendations »* (23).

En outre, le groupe a aussi examiné et accepté les résultats des 2 revues systématiques qu'il avait commissionnées. Ce travail leur a permis de publier dans la même déclaration de consensus en 2008, des recommandations sur l'utilisation des mesures d'intensité de douleur qui devraient être utilisées dans les études chez les enfants.

6.2.3 Critères de qualité des revues systématiques et des scores selon PedIMMPACT

6.2.3.1 Les degrés de preuve d'efficacité

Les auteurs des revues systématiques ont utilisé le système de notation suivant pour évaluer le niveau de preuve de l'efficacité des différentes échelles de la douleur (25):

- Efficacité « bien établie » (Niveau 1)
 - L'échelle a été publiée dans au moins 2 différents articles revus par des pairs, par au moins 2 différents groupes de chercheurs.
 - L'échelle est présentée avec suffisamment de détails pour permettre son évaluation critique et sa réplication.
 - Sa validité et sa fiabilité sont bonnes et sont détaillées (par exemple avec statistiques) et publiées dans au moins un article revu par des pairs.
- Efficacité « approchant le bien établi » (Niveau 2)
 - L'échelle a été publiée dans au moins 2 différents articles revus par des pairs, potentiellement par le même groupe de chercheurs.
 - L'échelle est présentée avec suffisamment de détails pour permettre son évaluation critique et sa réplication.
 - La validité et la fiabilité de l'échelle sont publiées, mais sans détails (par exemple sans statistiques) ou la validité et la fiabilité sont moyennes.
- Efficacité « avec bilan prometteur » (Niveau 3)
 - L'échelle a été publiée dans au moins 1 article revu par des pairs.
 - L'échelle est présentée avec suffisamment de détails pour permettre son évaluation critique et sa réplication.
 - La validité et la fiabilité sont publiées, mais sans détails (par exemple sans statistiques) ou la validité et la fiabilité sont moyennes.

6.2.3.2 Fiabilité, validité : définitions

La fiabilité d'une échelle est sa capacité à reproduire une mesure au fil de temps.

Les déterminants de la fiabilité sont (24):

1. Le « *inter rater reliability* » qui représente le degré d'accord entre plusieurs testeurs.
2. Le « *test-retest reliability* » qui représente le degré d'accord entre la même observation, effectuée sur la même personne à plusieurs reprises.
3. Le « *internal consistency* » qui représente la corrélation entre différents items dans le même test. Par exemple, si plusieurs items mesurent des éléments qui sont liés (par exemple : « pouvez-vous faire les courses? » et « pouvez-vous faire le ménage? »), et qu'ils varient dans le même sens, avec la même amplitude, cela sera indicateur d'une bonne « *internal consistency* » du test (26).

La validité d'une échelle est une mesure de sa capacité à mesurer ce qu'elle est sensée mesurer. Les déterminants de la validité sont (24) :

1. Le « *face validity* », qui représente le fait que l'échelle contient les éléments qui semblent appropriés pour ce qu'elle propose de tester.
2. Le « *content validity* » : les éléments de l'échelle comportent l'information appropriée.
3. Le « *criterion* » : la nouvelle échelle a été corrélée avec une échelle considérée comme le « *gold standard* ».
4. Le « *construct validity* » qui détermine la validité de variables abstraites qui ne peuvent pas être observées directement, telles que la douleur (26).

6.2.4 Domaines principaux de résultats

Le groupe PedIMMPACT a défini 6 domaines principaux de résultats qui devraient être considérés par les chercheurs dans les études cliniques de la douleur chez l'enfant.

Ces 6 principaux domaines de résultats sont :

1. L'intensité de la douleur (c'est ce domaine sur lequel cette thèse va se concentrer).
 - Identifier et évaluer la douleur.
2. La satisfaction globale des patients vis-à-vis du traitement.
 - Reflet de l'expérience des patients sous traitement.
3. Les symptômes et effets secondaires du traitement.
 - Tous les signes, symptômes, résultats de laboratoire ou maladies survenant après le début du traitement.
4. La récupération physique.
 - Aspects du fonctionnement physique du patient qui sont touchés par l'opération ou la procédure. Par exemple, temps requis avant de remarcher ou avant de recommencer à manger.
5. La réponse émotionnelle à la douleur.
 - Aspects de l'affect négatifs ou de la détresse secondaire à la douleur. Par exemple, l'anxiété ou la peur.
6. Les facteurs économiques.
 - Aspect financier qui peut être direct, comme le coût d'un médicament, ou indirect comme le coût du temps pendant lequel le parent ne peut pas travailler.

6.2.5 Les recommandations de mesures de la douleur

Le groupe a publié ses recommandations de mesures de la douleur dans les études chez les enfants. Un résumé de ces recommandations se trouve dans le tableau 1.

En général, les échelles d'auto-évaluation sont le premier choix pour les enfants capables de les utiliser. Les échelles comportementales sont nécessaires dans le cas d'enfants très jeunes et les enfants avec un retard de développement.

6.2.5.1 Les échelles d'auto-évaluation

Les 3 échelles d'auto-évaluation conseillées par le groupe PedIMMPACT sont :

- Le *Poker Chip Tool* (3 à 4 ans).

- Le *Faces Pain Scale* (4 à 12 ans).
- Le *Visual Analog Scale* (8 et plus ans).

6.2.5.1.1 Poker Chip Tool

Le *Poker Chip Tool* est une échelle de 0 à 4 qui est aussi connue sous le nom *Pieces of Hurt* ou « jetons de douleur ». Il y a 4 jetons qui représentent l'intensité de la douleur variant de « un peu » jusqu'à « la plus grande douleur imaginable » (27). Le soignant dit à l'enfant « Je veux parler avec toi à propos de la douleur que tu peux ressentir en ce moment. As-tu mal maintenant ? ». Si l'enfant répond non, un 0 de douleur est enregistré. Si l'enfant dit oui, on lui donne 4 jetons. On dit ensuite à l'enfant « ce sont des jetons de douleur – un jeton représente « une petite douleur », et 4 jetons « la plus grande douleur que tu puisses imaginer avoir ». La douleur que tu te ressens maintenant, à combien de jetons équivaut-elle ? 1, 2, 3 ou 4 ? » (28).

6.2.5.1.2 Faces Pain Scale Revised

Le *Faces Pain Scale Revised* (FPSR) consiste en une suite de 6 visages (figure 1) illustrant d'un côté le fait de ne ressentir « aucune douleur » jusqu'à « la plus grande douleur possible ». On dit à l'enfant « ces visages montrent combien on peut avoir mal » et puis on explique que « le visage à gauche montre quelqu'un qui n'a pas mal du tout » et les visages suivants montrent « quelqu'un qui a de plus en plus mal » jusqu'au le dernier « qui montre quelqu'un qui a très très mal ». Finalement, on demande à l'enfant d'indiquer « le visage qui montre combien tu as mal en ce moment ». The FPSR est ensuite converti en une échelle de 0 à 10 où le premier visage vaut 0, le deuxième 2, le troisième 4, le quatrième 6, le cinquième 8 et le dernier 10 (24, 29, 30).

6.2.5.1.3 Visual Analog Scale

Le *Visual Analog Scale* (VAS) consiste en une ligne horizontale, ou verticale, avec les limites de lignes qui représentent les extrêmes de la douleur (figure 2). On demande à l'enfant de sélectionner ou marquer un point sur la ligne qui représente l'intensité de sa douleur, variant de « pas de douleur du tout » à gauche ou en bas,

à « la pire des douleurs imaginables » à droite ou en haut (23). Cette échelle peut varier de 0 à 10 (cm), ou de 0 à 100 (mm).

6.2.5.2 Les échelles comportementales

Les échelles comportementales conseillées par PedIMMPACT sont au nombre de 4 et comprennent :

- Le *Face, Legs, Activity, Crying, Consolability* (FLACC) (pour les enfants 1 an et plus).
- Le *Children's Hospital of Eastern Ontario Pain Scale* (CHEOPS) (1 an et plus).
- Le *Toddler-Preschooler Postoperative Pain Scale* (TPPPS) (1 à 5 ans).
- Le *Parents' Postoperative Pain Measure* (PPPM) (1 an et plus, pour être utilisé à la maison par les parents).

A noter que, même si ces échelles sont conseillées pour être utilisées par des enfants à partir de 1 an selon PedIMMPACT, la littérature suggère que les échelles CHEOPS et FLACC peuvent être utilisées à partir de 1 mois déjà (19).

6.2.5.2.1 FLACC

L'échelle *Faces, Legs, Activity, Cry, Consolability* (FLACC) comprend 5 catégories, et chaque catégorie comptabilise entre 0 et 2 points, résultant en une échelle de 0 à 10 points. Les 5 catégories décrivent l'expression du visage, la position de jambes, l'agitation du corps, si l'enfant pleure, et s'il est consolable (tableau 2) (31). Bien que toutes les mesures soient recommandées pour la douleur postopératoire, le FLACC ayant été conçu et validé dans ce contexte sur une tranche d'âge plus large, cette échelle est donc recommandée comme premier choix pour le post opératoire en milieu hospitalier (19).

6.2.5.2.2 CHEOPS

L'échelle CHEOPS comprend 6 catégories : la présence de pleurs (score 1 à 3), l'expression faciale (0 à 2), l'expression verbale (0 à 2), la position du torse (1 à 2), la position des mains (1 à 2), et la position des jambes (1 à 2). Puisque 4 catégories ont un score minimum de 1, le score minimum pour le CHEOPS est 4. Cinq

catégories ont un maximum de 2 et une catégorie a un maximum de 3 donc le score maximale est 13. Alors l'échelle est de 4 à 13 points (tableau 3) (5).

6.2.5.2.3 TPPPS

L'échelle *Toddler Preschooler Postoperative Pain Scale* (TPPPS) comprend 7 points dans 3 catégories : expression verbale - pleur, gémissement et cri (0 à 3 points), expression faciale - bouche ouvert, yeux fermés et renflement de sourcils (0 à 3 points), et expression corporelle (1 point). L' échelle varie donc de 0 à 7 (pour plus de détail voir tableau 4) (32).

6.2.5.2.4 PPPM

L'échelle *Parents' Postoperative Pain Measure* (PPPM) est une échelle utilisée par les parents, pour quantifier la douleur de leur enfant, à la maison. Elle comprend 15 catégories et un score total variant de 0 à 15 points. Les catégories sont, par exemple, si l'enfant pleure, joue comme d'habitude, mange, mais également si les parents sont inquiets (tableau 5) (33).

6.2.6 PedIMMPACT et l'utilisation des échelles de la douleur dans la recherche

Dix ans après la publication des recommandations du groupe PedIMMPACT, peu de revues systématique chez les enfants, semblent prendre en considération les domaines de résultats et mesures conseillés par ces recommandations (34). De plus, il semblerait que seul un tiers des auteurs de revues systématiques et d'études randomisées contrôlées se focalisant sur la douleur post-opératoire chez l'enfant connaissent même la simple existence de ces domaines de résultats conseillés par PedIMMPACT. Les raisons pour lesquelles ces auteurs n'utilisent pas les échelles recommandées restent en majeure partie inconnues (35). Les raisons du choix d'une échelle de douleur donnée comme outil d'évaluation de l'efficacité thérapeutique d'un médicament ou d'une procédure lors d'une conception d'un nouvel essai clinique randomisé et contrôlé, sont inconnues. C'est à cette question que cette thèse va tenter de répondre.

7 Objectifs de ce travail

L'objectif primaire de cette thèse est de mieux comprendre pourquoi, et sur quelles bases, les chercheurs choisissent une échelle de douleur donnée plutôt qu'une autre pour mesurer la douleur dans leurs études randomisées contrôlées effectuées en période péri-opératoire chez l'enfant. Dans le cadre de ce travail, et afin d'étudier un groupe le plus homogène possible, je me suis concentrée sur les études testant un seul médicament, le tramadol, qui est de plus en plus utilisé chez l'enfant depuis l'émergence de restrictions à l'utilisation de la codéine.

L'objectif secondaire de ce travail est de vérifier si ces études randomisées publiées utilisent les échelles qui sont conseillées par PedIMMPACT. L'évaluation de l'influence de la publication des recommandations établies par le groupe PedIMMPACT sur les choix des échelles de la douleur dans les études cliniques sera effectuée en recherchant une augmentation éventuelle de l'utilisation des échelles de douleurs conseillées par PedIMMPACT après 2008.

8 Methods

8.1 Study design

This study was designed as a cross-sectional analysis of systematically searched pediatric trials reporting on postoperative pain. In order to obtain a homogenous set of trials, we focused on randomized controlled trials (RCT) testing one single drug, tramadol, for the control of post-operative pain.

8.2 Setting

8.2.1 Trials and search strategy

We systematically searched for all published randomized controlled trials testing tramadol for the prevention or treatment of post-operative pain in children, published between 2000 and 2017 inclusively (9 years prior to, and 9 years after the publication of the PedIMMPACT guidelines).

We searched in Medline (via Pubmed), the Cochrane Library and Embase up until January 2018. We also hand searched bibliographies of relevant review articles and systematic reviews.

There were no date or language restrictions imposed. In Medline and the Cochrane Library we searched for “[tramadol] or [tramal]” and “[ped] or [paed] or [child]”. In Embase, the search criteria were “[tramal] or [tramadol] and “[randomised control trial] or [systematic review] or [meta-analysis]” and “[preschool] [school] [child] [adolescent]”.

8.2.2 Inclusion criteria

RCTs in children (1 month to 16 years) undergoing surgery under general anesthetic, testing tramadol versus active treatment or placebo, which reported on post-operative pain were included. Tramadol could be given preoperatively (premedication), intraoperatively or postoperatively via any mode of administration.

8.2.3 Exclusion criteria

We did not consider non-RCTs, dose response studies without control group, studies with a multi drug systemic treatment group, procedures under local anesthetic only and trials in neonates or adults (>16 years).

8.3 Variables

From each trial, the following variables were extracted information on: 1st author, journal, year and language of publication, trial information including: country performed in, surgery type, mode of administration of tramadol and control drug, sample size and calculation, funding (industry, academic, none, not declared), protocol registration, risk of bias, and age range of the children analyzed.

The pain intensity scale used in the RCT was extracted, including any justification provided for the choice. We extracted references cited when stating or discussing the pain score used, when present, and particularly whether a reference to the PedIMMPACT study was present (yes/no). Where there was not an explanation regarding choice of pain score in the article, the corresponding author was contacted. Authors were asked which factors aided them in choosing the pain score(s) used in their study. They were able to respond in free text and examples were also given: current pain score used in hospital, prior published articles, systematic review or guidelines.

8.4 Data sources/measurement

The full text of all included articles was searched. We extracted the pain intensity scale used from the description in the text, figures and references. Where scales had different names but used the same scoring system they were classified as the same intensity score. For example, Pain Discomfort Scale is also known as the Modified Hannallah Pain Score.

If the same pain intensity score was used in different articles, but different scales applied to the score, they were considered to be 2 different scores. For example, CHEOPS 13 (scale 3 to 14) and modified CHEOPS (scale 0 to 10) were considered to

be 2 different pain scales. The visual analogue scale (VAS) using 100 mm scale was considered to be similar to the scale of 0 to 10 cm. Where the scale of the pain score was not described or unclear, it was listed as a separate pain intensity score (undefined scale).

For the descriptions of control drugs used, normal saline, physiological solution and placebo were considered similar.

Authors were initially contacted via the published email for correspondence. Where this failed to be delivered, the internet was searched for alternative email addresses based on the names and affiliations of the authors.

8.5 Bias

Inclusion of studies was performed by 1 author* and where there was uncertainty clarified with a co-author. One author† extracted the data. This was independently checked by a co-author and any discrepancies resolved with a further author.

8.6 Study size

There was no pre-hoc hypothesis on any association to be tested. Therefore, the study size was chosen to gather an exhaustive sample of RCTs testing tramadol in children with approximately an equal number of them published before and after the publication of the PedIMPACT recommendations.

8.7 Quantitative variables

The reasons given for the choice of pain scores used was firstly recorded as described in the original report.

These reasons were then categorized, by us, into 4 main categories:

* L'auteur de cette thèse

† L'auteur de cette thèse

- 1) « Validated » pain scale: when authors justified their choice of a specific pain scale by the fact that this pain scale was validated (with or without references).
- 2) « Popular » pain scale: when authors justified their choice saying that it was a scale “used by others” or described as “popular” (with or without references)
- 3) « Familiar » pain scale: when authors justified their choice saying that it was commonly used or a standard in their hospital
- 4) « PedIMMPACT » when their choice was based on PedIMMPACT recommendations.

Some authors of trials provided references to support their choice of a specific pain score. These references were extracted, and categorized, by us, into 6 main categories:

- 1) « Validating » reference: reference of a study whose main objective was to validate the pain scale used.
- 2) « Prior study »: reference of a study that previously used that pain scale.
- 3) « Review article »: reference to a pediatric pain review article, chapter or systematic review of pediatric pain scales.
- 4) « Ease of use »: simple, accessible and easy implementation.
- 5) « PedIMMPACT »: reference to PedIMMPACT publication.
- 6) « Out of context »: reference to a study that does not focus on the pain scale used (for example, that focuses on a different pain scale or in a different context).

The responses from authors were classified using the above definitions as well as:

- 1) « Other »: not fitting any of the above classifications

8.8 Statistical methods

Results are presented as numbers (proportions), means (standard deviations) or median (interquartile range) as required based on the underlying distribution of the

data, and according to the period of time (year) to examine a potential difference due to the publication of PedIMPACT.

9 Results

9.1 Trial selection

483 articles were initially identified and 373 excluded due to being out of context. Of the remaining 110 articles, 45 were excluded due to not being a randomized controlled trial (18), procedures not performed under general anesthetic (8), having no control group (6), not being full articles, or us being unable to access full article despite contacting journal directly (6), being in neonates (4), not reporting on pain outcomes (1), using systemic multi drug regimen (1) and not testing tramadol (1). This resulted in 65 eligible studies. A further 6 studies were written in languages that we were unable to translate (Turkish (2), Slovenian, Chinese, Korean, Spanish (1 each)) thus were excluded. Fifty-nine RCTs remained to be analyzed (figure 3).

9.2 Descriptive data

Trials were performed in 15 countries. These were Turkey (26 trials), India (8), Iran (6), Pakistan (4), Egypt (3), Taiwan and Finland (2 each), Australia, Belgium, Brazil, Italy, Mexico, Nepal, Scotland, and USA (1 each).

Fifty-four (92%) trials focused on Ear Nose and Throat (ENT) or infra-umbilical surgery.

Tramadol was given intravenously (21 trials), via caudal (20), infiltration (7), epidural (3), topically (2), orally (1), intramuscularly (1), both infiltration and intramuscularly (2), both topically and intravenously (1) and the route of administration was unstated in 1 trial.

The 59 RCTs included 70 control arms (multiple controls used in some trials). The drugs used in the control arms were local anesthetics with or without adrenaline (24 trials), normal saline (16), morphine (7), ketamine (6), pethidine (3), paracetamol (3), dexmedetomidine (2), fentanyl (2), nalbuphine (2) codeine (1), midazolam (1), neostigmine (1), ketorolac (1) and nil (1).

Trial sizes ranged from 20 to 150 (median 60, interquartile range 45-80). Sample size calculation was reported in 27 (46%) of trials. The ages of children examined ranged from 4 months to 16 years.

Nine trials (15%) declared sponsorship, of which 7 were academic funding (from a hospital or university research fund) and 2 were from pharmaceutical companies. Seven trials (12%) stated they had no funding and the remaining 43 (73%) did not make any declaration of funding.

Three trials (5%) had registered their protocol in a trial registry (1 at ClinicalTrials.gov and 2 at Iranian registry of clinical trials). However, the latter 2 were retrospectively registered.

Risk of random sequence bias was low in 59% of trials and high in 5%. The risk of selection bias was high in 18% of trials and unclear in 75%. The risk of performance bias was low in 46% of trials and unclear in 51%. Detection bias was low in 59% and unclear in 39%. Attrition bias was low in 25% and unclear in 71%. Reporting bias was low in 1 study and was unclear in the remaining 58 studies. See risk of bias tables for further detail (figure 4).

9.2.1 Pain scores used

There was a total of 39 different pain score used across all analyzed trials, of which 22 were observational pain scores, 15 were self-reported pain scales, and 2 were scales performed by parents at home (1 self-report, 1 not stated). 48 trials used 1 pain score each, 8 used 2 pain scores, and 3 trials used 3 pain scores. See table 6 for a full description of the pain scores used.

9.2.1.1 Observational pain scores

The 22 observational pain scores were used across 45 RCTs. These included 8 different versions of what was initially the Pain Discomfort Scale which has had small modifications subsequently and been renamed both the Modified Hannallah Pain Score and the Objective Pain Scale (used in 13 trials). There were 4 variations of the Children's Hospital of Eastern Ontario Pain Scale (CHEOPS) pain score (16

trials). There was a single version of the following pain scores: the Face, Legs, Activity, Cry, Consolability score (FLACC) (5 trials), the modified Toddler-Preschool Postoperative Pain Scale (TPPPS) (4 trials), the All India Institute of Medical Sciences (AIIMS) pain discomfort scale (2 trials), the Maunuksela pain score, a hemodynamic score, the Children's and Infants' Postoperative Pain Scale (CHIPPS scale), the four-point restlessness–pain protocol, the Prince Margaret Hospital pain score, a pain scale on swallowing, and a 3 point pain scale (used in 1 trial each).

9.2.1.2 Self-reported pain scores

The 15 self-reported pain scores were used across 20 RCTs. Seven different Faces pain scales were used in 13 trials. These included the Faces Pain Scale Revised (3 trials), the Faces Pain Scale 0 to 5 (2 trials), the Faces Pain Scale 0 to 6 (2 trials), the Wong Baker Faces pain scale 0 to 5 (3 trials), the Wong Baker Faces pain scale 0 to 10 (1 trial), the McGrath Facial Scale 0 to 8 (1 trial) and the McGrath's Face Scale 0.04 to 0.97 (1 trial).

The remaining 8 self-reported pain scales were used in 11 trials and included the Visual Analogue Scale (VAS) (4 trials), the Numerical Rating Scale (NRS) 0 to 5, the NRS-11 scale 0 to 10, the verbal rating scale, a simple descriptive scale, a 5-point scale, a verbal expression scale, and an unnamed pain score (1 trial each).

9.2.1.3 Pain scores used by parents

The 2 scales used by the parents were reported in 1 trial each and included a categorical pain scale, and an unspecified scale.

9.3 Outcome data

9.3.1 Primary outcomes

9.3.1.1 Justification of the pain score used

Reasons provided for the pain scale chosen varied across trials.

13 trials explained their choice of pain intensity score in the text of the article and included a reference, 31 included a reference but provided no explanation for their choice of pain score and 15 included neither an explanation nor a reference regarding the pain score used (figure 5). The authors of these later 2 groups (total 46 RCTs) were contacted regarding their choice of pain score.

Of the 13 trials which included an explanation of their choice of pain score, 4 (31%) chose the scale(s) because they were “validated” pain scales, 3 (23%) because they were “validated and familiar” pain scales, 2 (15%) because they were both “validated and popular” pain scales, 2 (15%) because they were “popular” pain scales, 1 (8%) because it was “validated, popular and easy to use” and 1 (8%) because it was recommended by PedIMMPACT.

Of the 46 authors contacted, 6 answered our inquiries. Two authors indicated that prior published studies and review articles aided them in choosing the pain score that they used. One author stated their choice was due to prior published review articles and familiarity with the pain score chosen. One author stated familiarity with the chosen pain score. A further author chose the pain study based upon ease of use particularly easy accessibility and implementation. A final author was unable to recall. See table 7 for a full description of responses.

9.3.1.2 References provided to justify the choice of the pain scale used

Forty-four trials included at least 1 reference to justify their choice of the pain scale used. In total there were 68 references. Thirty-three trials provided 1 reference, 2 provided 2 references, 6 provided 3 references, 2 provided 4 references and 1 provided 5 references.

Of the 68 references, 27 (40%) were references to prior studies using the same pain scale, 26 (38%) were references validating the pain scale, 8 (12%) were references to a review article, 1 (1%) referenced PedIMMPACT, and 6 (9%) were out of context.

9.3.2 Secondary outcomes

Seventeen (29%) trials analyzed used at least 1 pain scale which was recommended by the PedIMMPACT statement, whereas 42 (71%) did not.

There was no increase in the use of pain scores recommended by the PedIMMPACT statement after its publication. From 2000 to 2008, 9 of 30 RCTs (30%) had used at least 1 pain score which was recommended by PedIMMPACT versus 8 of 29 (28%) from 2009 to 2017 (figure 6).

Of the 17 studies with at least 1 pain scale recommended by PedIMMPACT, 13 (76%) used the recommended pain scale in the age group suggested by PedIMMPACT whereas 4 (24%) did not.

10 Discussion

10.1 Key results

Of the 59 trials analyzed, 13 included a written explanation of their choice of pain score including references to support their choice, a further 31 provided a reference only, without explanations, while authors of 15 trials provided neither an explanation nor a reference to justify their choice. The explanations for author choice of pain score where provided in the trial were most commonly that the score was validated (31%) or validated and familiar (23%). References when provided, were most commonly to a prior article using the same pain score (40%) or to an article validating the score (38%). Of all the trials analyzed, only 1 discussed or referenced the PedIMMPACT statement. Finally, the number of trials using pain score recommended by the PedIMMPACT statement was low and did not increase after its publication.

10.2 Interpretation

The ability to compare the results of pain studies in children and adolescents has been limited due to the variety and number of pain intensity scores used. A Cochrane review of tramadol for post-operative pain treatment in children highlighted methodological problems particularly the use of different validated and non-validated pain scales in the studies. This resulted in difficulty ascertaining the efficacy of tramadol for the post-operative treatment of pain in children (36). Equally, a systematic review of pain relief after tonsillectomy found the varied methodology including methods of pain measurement resulted in difficulty drawing conclusions as to the clinical efficacy of the analgesia regimes being investigated. The review found time to rescue analgesia and the need for rescue analgesia to be more useful than pain intensity (37). The PedIMMPACT statement attempted to address this problem.

The lack of ability to perform a meta-analysis due to the variety of outcomes reported is not only an issue in pediatric pain systemic reviews, but also in pediatric reviews not investigating pain (38).

Our study highlights that there are a great variety of pain scales used in pediatric trials, even when we limit the field to 1 specific drug, such as tramadol. Moreover, it seems that there has been no change in the use of pain intensity scales since the PedIMMPACT statement. The relative limited impact of published systematic reviews on subsequent trials has already been shown in the field of anesthesiology (39) and the reason for this is likely to be multifactorial.

The lack of implementation of methodological recommendations in clinical trials is not unique to pediatrics and has been identified in the adult population also. Barriers to the implementation of clinical guidelines include knowledge factors such as lack of knowledge or familiarity, attitude factors such as lack of agreement, guideline factors such as lack of evidence and external factors such as lack of resources (40, 41).

Lack of knowledge of the existence of the PedIMMPACT guidelines is likely to be a significant factor in our study population. A recent article suggested that only one third of respondents of a survey had heard of the 6 recommended core impact domains. Steps to increase awareness of the PedIMMPACT guidelines such as education may increase the use of their recommended outcomes and measures. The same study also identified the complexity of the guidelines, and lack of resources available to researchers contributed to non-compliance with the PedIMMPACT statement (35).

Geographical factors may also contribute. Guidelines are commonly developed by national and international guideline development groups (42). As a result, there is commonly variation in evidence based practice internationally (43). The PedIMMPACT statement was developed by a North American group. The only trial in our study to reference the guidelines was performed in the United States. All the other studies were performed outside of the United States. This was recognized to be a factor in reporting of core outcome domains in pain studies in adults (44).

Given the number of pain intensity scores used in pediatric studies, the fact that less than a quarter of authors described their choice of pain score is concerning. In addition, in some studies the actual pain intensity score used or the scale applied was not clear. This has also previously been reported by authors attempting to perform a meta-analysis of pediatric analgesia post tonsillectomy (37).

The relevance of references related to pain scores used varied in the trials we analyzed. For example, in 4 trials, the references related to different pain scales to those used in their own trial. Prior analysis of articles in published medical journals have identified quotation errors in approximately 15% of references (45, 46). This is a higher rate to what we identified (6 of 68 references i.e. 9%) however we only analyzed pain references.

10.3 Strengths and limitations

Our study has strengths but also some limitations.

The major strength of our work is that we focused on a single drug, and a single study design in a limited population. The reason for this choice was due to increasing use of tramadol in pediatric patients as a result of restrictions in the use of codeine, despite some restrictions from the US Food and Drug Administration (FDA) for its administration (47). Despite this relatively homogeneous selection, the variability of the pain scale used remained substantial. We performed a systematic search for all published trials fulfilling our pre-defined inclusion criteria without language restriction. Although we may have missed a trial, our strategy enabled us to analyze trials performed in different parts of the world.

Our study also suffers some limitations. Firstly, when several versions of a pain scale exist, we only considered the exact version recommended by the PedIMMPACT statement. For example, CHEOPS 4-13, the original version, is recommended by PedIMMPACT. A modified CHEOPS score has subsequently been developed from 0 to 10 to make it easier to compare results (48) but it is not the version recommended by PedIMMPACT. Similarly, the TPPPS is recommended, but the modified version is not. Although this strict interpretation can be questioned,

this is in line with the observational pain scale systematic review commissioned by PedIMMPACT which chose not to include modified versions as they make it difficult to compare the results to the original published measure (19). On the other hand, PedIMMPACT does recommend the revised version of the Faces Pain Scale which was developed partly as it uses a scale of 0 to 10 versus the original scale of 0 to 6 (23). The systematic review of self-report pain scales systematic review commissioned by PedIMMPACT recommends this scale as it is “the most psychometrically sound measure in school-aged children” (24). This difference in approach is not explained in the PedIMMPACT statement, however, it may be due to different approaches by the different authors for each systematic review. To maintain consistency, we chose to only include the exact version as recommended by the PedIMMPACT statement. Some authors used sections of a validated pain score, for example, using only 2 categories of the modified CHEOPS with a scale of 0 to 4. We equally did not consider this to be a scale recommended by PedIMMPACT.

Second, since the publication of the PedIMMPACT statement, further research has occurred on the same topic. A systematic review of face scales, published in 2010, reported that children preferred the Wong Baker Faces pain rating scale to the Faces Pain Scale, Faces Pain Scale Revised and Oucher pain scale (49). One study listed this as the reason for choosing this scale. A study validating the use of the Numerical Rating Scale in children and adolescents has also been published since the PedIMMPACT statement (50). This evidence would be considered in any revised version of the PedIMMPACT but was not considered in our study.

Third, the PedIMMPACT group aimed to address studies in children from 3 to 16 years. However, where their recommended pain intensity scales have been validated in younger children they have included this information in their recommendations. For example, all the behavioral pain scales have been validated for children from 1 year of age. We chose to include trials down to but excluding neonates in order to broaden the criteria and include trials which had both younger and older children. All trials included in our study had at least some children who were 3 years or older.

Fourth, the PedIMMPACT guidelines were published in 2008 however it is possible some RCTs published in 2009 and onwards were commenced prior to the PedIMMPACT publication.

Finally, the response rate from authors was low with only 13% of authors responding. This makes generalizations about their responses difficult to interpret.

10.4 Generalizability

The choice of pain intensity score in pediatric research studies continues to vary widely and the ability to perform a meta-analysis continues to be limited. The work by the PedIMMPACT group is a significant step towards uniformity in pediatric pain studies, however, this has not translated into changes in published studies. The development of core outcome domains includes periodic review and the results of this research should be considered in any such review. A focus on education of researchers and increased awareness of the PedIMMPACT statement may increase the use of the pain intensity domains recommended by the group.

11 Bibliography

1. Société Française d'Etude et de Traitement de la Douleur. Définition <http://www.sfetd-douleur.org/definition> [Accessed 10 March 2019].
2. Huguet A, Stinson JN, McGrath PJ. Measurement of self-reported pain intensity in children and adolescents. *J Psychosom Res.* 2010;68(4):329-36.
3. Baker DW. Joint Commission Statement on Pain Management https://www.jointcommission.org/joint_commission_statement_on_pain_management/ [Accessed 10 March 2019].
4. Cousins MJ, Brennan F, Carr DB. Pain relief: a universal human right. *Pain.* 2004;112(1-2):1-4.
5. McGrath P, Johnson G, Goodman J, McGrath P, Johnson G, Schillinger J, et al. CHEOPS: a behavioral scale for rating postoperative pain in children. *Adv Pain Res Ther*; 1985; 9:395-402.
6. Eland JM, Anderson JE. The experience of pain in children. A source book for nurses and Other Health Professionals: Little, Brown and Company 1977.
7. Young KD. Pediatric procedural pain. *Ann Emerg Med.* 2005;45(2):160-71.
8. Dowden S, McCarthy M, Chalkiadis G. Achieving organizational change in pediatric pain management. *Pain Res Manag.* 2008;13(4):321-6.
9. Vittinghoff M, Lönnqvist PA, Mossetti V, Heschl S, Simic D, Colovic V, et al. Postoperative pain management in children: Guidance from the pain committee of the European Society for Paediatric Anaesthesiology (ESPA Pain Management Ladder Initiative). *Paediatr Anaesth.* 2018;28(6):493-506.
10. Perquin CW, Hazebroek-Kampschreur AA, Hunfeld JA, Bohnen AM, van Suijlekom-Smit LW, Passchier J, et al. Pain in children and adolescents: a common experience. *Pain.* 2000;87(1):51-8.
11. Cohen LL, Lemanek K, Blount RL, Dahlquist LM, Lim CS, Palermo TM, et al. Evidence-based assessment of pediatric pain. *J Pediatr Psychol.* 2008;33(9):939-55; discussion 56-7.
12. Treadwell MJ, Franck LS, Vichinsky E. Using quality improvement strategies to enhance pediatric pain assessment. *Int J Qual Health Care.* 2002;14(1):39-47.
13. Beltramini A, Milojevic K, Pateron D. Pain Assessment in Newborns, Infants, and Children. *Pediatr Ann.* 2017;46(10):e387-e95.
14. von Baeyer CL. Children's self-reports of pain intensity: scale selection, limitations and interpretation. *Pain Res Manag.* 2006;11(3):157-62.
15. Pédiadol. Evaluation de la Douleur Chez l'Enfant <https://pediadol.org/evaluation/> [Accessed 13 March 2019].
16. von Baeyer CL. Children's self-report of pain intensity: what we know, where we are headed. *Pain Res Manag.* 2009;14(1):39-45.
17. Champion GD, Goodenough B, von Baeyer CL, Thomas W. Measurement of pain by self-report. *Measurement of Pain in Infants and Children.* 10. Seattle: IASP Press; 1998. p. 123-60.
18. von Baeyer CL, Spagrud LJ. Systematic review of observational (behavioral) measures of pain for children and adolescents aged 3 to 18 years. *Pain.* 2007;127(1-2):140-50.

19. The Centre for Evidence-Based Medicine. Oxford Centre for Evidence-based Medicine – Levels of Evidence <https://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/> [Accessed 12 March 2019].
20. McGrath PJ, Walco GA, Turk DC, Dworkin RH, Brown MT, Davidson K, et al. Core outcome domains and measures for pediatric acute and chronic/recurrent pain clinical trials: PedIMMPACT recommendations. *J Pain*. 2008;9(9):771-83.
21. Stinson JN, Kavanagh T, Yamada J, Gill N, Stevens B. Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents. *Pain*. 2006;125(1-2):143-57.
22. Cohen LL, La Greca AM, Blount RL, Kazak AE, Holmbeck GN, Lemanek KL. Introduction to special issue: Evidence-based assessment in pediatric psychology. *J Pediatr Psychol*. 2008;33(9):911-5.
23. Streiner D, Norman G, Cairney J. *Health Measurement Scales A practical guide to their development and use*. 5th Edition ed. New York: Oxford University Press; 2015.
24. Hester N, Foster R, Kristensen K. Measurement of pain in children: generalizability and validity of the pain ladder and Pieces of Hurt tool. *Adv Pain Res Ther*. 1990(15):79–84.
25. Hester NK. The preoperational child's reaction to immunization. *Nurs Res*. 1979;28(4):250-5.
26. Hicks CL, von Baeyer CL, Spafford PA, van Korlaar I, Goodenough B. The Faces Pain Scale-Revised: toward a common metric in pediatric pain measurement. *Pain*. 2001;93(2):173-83.
27. International Association for the Study of Pain. FPS-R French Instructions <https://www.iasp-pain.org/Education/Content.aspx?ItemNumber=1823&navItemNumber=1119>: International Association for the Study of Pain; [Accessed 3 April 2019].
28. Merkel SI, Voepel-Lewis T, Shayevitz JR, Malviya S. The FLACC: a behavioral scale for scoring postoperative pain in young children. *Pediatr Nurs*. 1997;23(3):293-7.
29. Tarbell SE, Cohen IT, Marsh JL. TPPPS (Toddler Preschooler Postoperative Pain Scale). *Pain*; 1992; 50: 273-80.
30. Chambers CT, Reid GJ, McGrath PJ, Finley GA. Development and preliminary validation of a postoperative pain measure for parents. *Pain*. 1996;68(2-3):307-13.
31. Boric K, Dosenovic S, Jelcic Kadic A, Boric M, Jeric M, Puljak L. Efficacy and Safety Outcomes in Systematic Reviews of Interventions for Postoperative Pain in Children: Comparison Against the Recommended Core Outcome Set. *Pain Med*. 2018;19(11):2316-21.
32. Boric K, Boric M, Dosenovic S, Jelcic Kadic A, Batinic M, Cavar M, et al. Authors' lack of awareness and use of core outcome set on postoperative pain in children is hindering comparative effectiveness research. *J Comp Eff Res*. 2018;7(5):463-70.
33. Schnabel A, Reichl SU, Meyer-Frießem C, Zahn PK, Pogatzki-Zahn E. Tramadol for postoperative pain treatment in children. *Cochrane Database Syst Rev*. 2015(3):CD009574.

34. Hamunen K, Kontinen V. Systematic review on analgesics given for pain following tonsillectomy in children. *Pain*. 2005;117(1-2):40-50.
35. Summerbell CD, Moore HJ, Vögele C, Kreichauf S, Wildgruber A, Manios Y, et al. Evidence-based recommendations for the development of obesity prevention programs targeted at preschool children. *Obes Rev*. 2012;13 Suppl 1:129-32.
36. Habre C, Tramèr MR, Pöpping DM, Elia N. Ability of a meta-analysis to prevent redundant research: systematic review of studies on pain from propofol injection. *BMJ*. 2014;348:g5219.
37. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA*. 1999;282(15):1458-65.
38. Fischer F, Lange K, Klose K, Greiner W, Kraemer A. Barriers and Strategies in Guideline Implementation-A Scoping Review. *Healthcare (Basel)*. 2016;4: 36-51.
39. Kredo T, Bernhardtsson S, Machingaidze S, Young T, Louw Q, Ochodo E, et al. Guide to clinical practice guidelines: the current state of play. *Int J Qual Health Care*. 2016;28(1):122-8.
40. Kramer JM, Newby LK, Chang WC, Simes RJ, Van de Werf F, Granger CB, et al. International variation in the use of evidence-based medicines for acute coronary syndromes. *Eur Heart J*. 2003;24(23):2133-41.
41. Mulla SM, Maqbool A, Sivananthan L, Lopes LC, Schandelmaier S, Kamaleldin M, et al. Reporting of IMMPACT-recommended core outcome domains among trials assessing opioids for chronic non-cancer pain. *Pain*. 2015;156(9):1615-9.
42. de Lacey G, Record C, Wade J. How accurate are quotations and references in medical journals? *Br Med J (Clin Res Ed)*. 1985;291(6499):884-6.
43. Mogull SA. Accuracy of cited "facts" in medical research articles: A review of study methodology and recalculation of quotation error rate. *PLoS One*. 2017;12(9):e0184727.
44. Rodieux F, Vutskits L, Posfay-Barbe KM, Habre W, Piguet V, Desmeules JA, et al. When the Safe Alternative Is Not That Safe: Tramadol Prescribing in Children. *Front Pharmacol*. 2018;9:148.
45. Splinter W, Semelhago L, Chou S. The Reliability and Validity of a Modified CHEOPS Pain Score. *ANESTH ANALG* 1994;78:S1-S503).
46. Tomlinson D, von Baeyer CL, Stinson JN, Sung L. A systematic review of faces scales for the self-report of pain intensity in children. *Pediatrics*. 2010;126(5):e1168-98.
47. von Baeyer CL, Spagrud LJ, McCormick JC, Choo E, Neville K, Connelly MA. Three new datasets supporting use of the Numerical Rating Scale (NRS-11) for children's self-reports of pain intensity. *Pain*. 2009;143(3):223-7.

12 Figures

Figure 1. The images used in the Faces Pain Scale – Revised.

Figure 2. Example of the Visual Analog Scale

Figure 3. Flow diagram of included/excluded studies

Figure 4. Risk of bias of included studies

Figure 5. Summary of choice of pain score used in trials

Figure 6. Summary of trials complying with PedIMMPACT per year

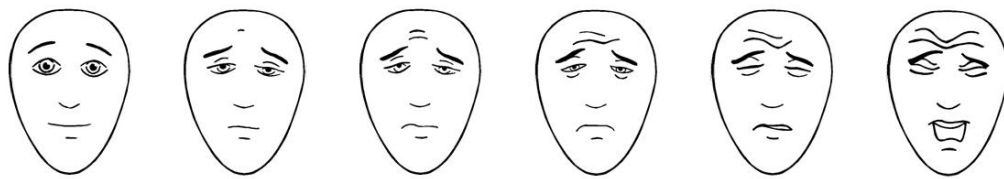


Figure 1. The images used in the Faces Pain Scale – Revised

Scale 0, 2, 4, 6, 8 and 10.

©2001, International Association for the Study of Pain.

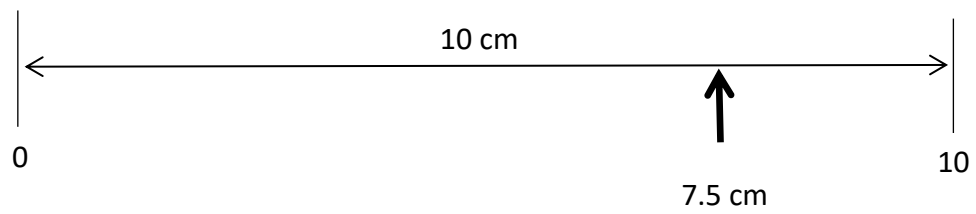


Figure 2. Example of the Visual Analog Scale

The arrow represents the point where the child marked the line. In this case, the result which is measured with a ruler would be 7.5.

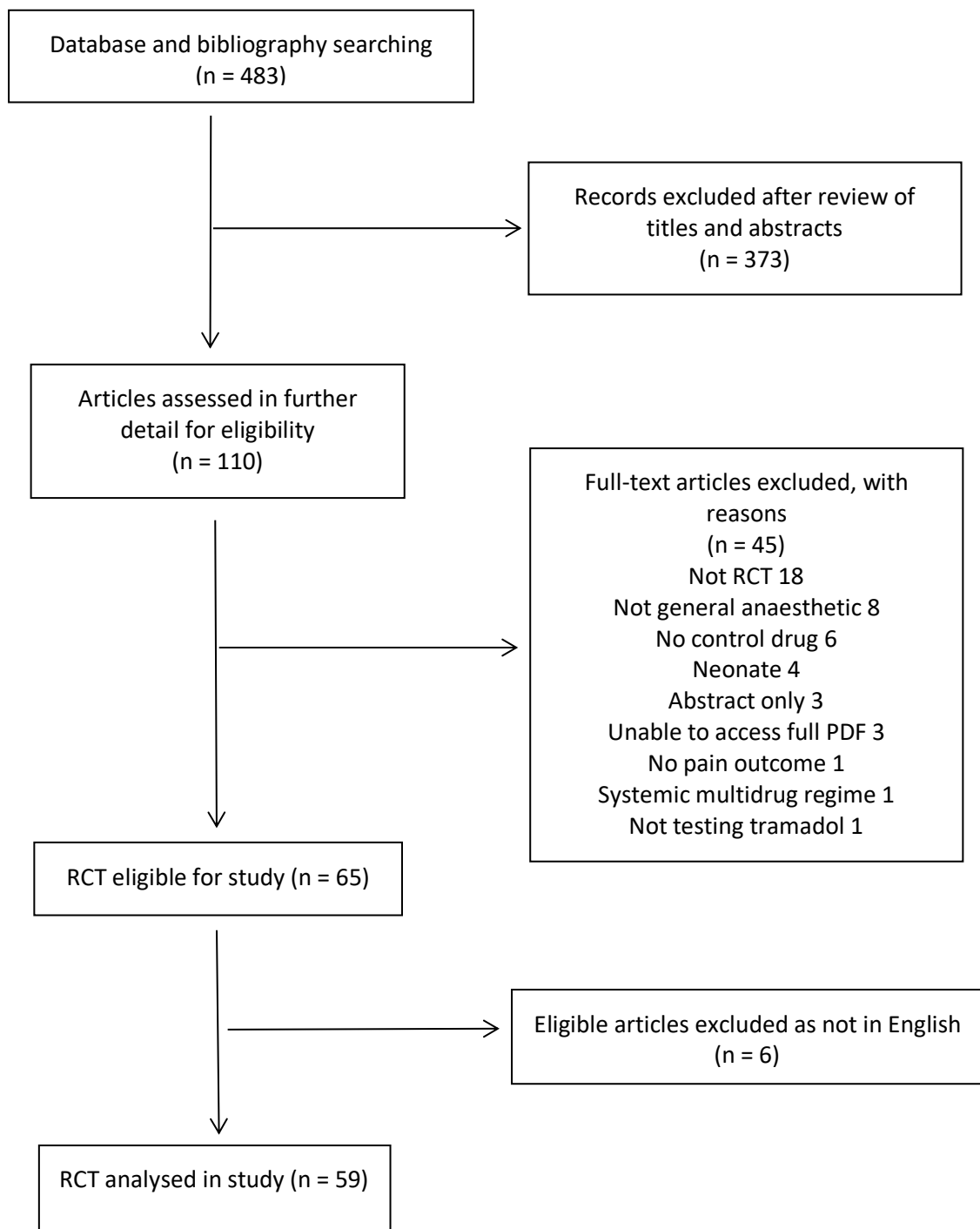


Figure 3. Flow diagram of included/excluded studies

Description of the number of articles identified with search criteria and reasons for exclusion of studies.

	Random sequence generation? (selection bias)	Allocation concealment? (selection bias)	Blinding of participants and personnel? (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data? (attrition bias)	Selective reporting? (reporting bias)
Akbay 2010	?	?	?	?	?	?
Ali 2008	?	?	?	?	?	?
Antila 2006	?	?	?	?	?	?
Ataf 2008	?	?	?	?	?	?
Ayatollahi 2012	?	?	?	?	?	?
Aydin 2013	?	?	?	?	?	?
Bedirli 2017	?	?	?	?	?	?
Chiaretti 2000	?	?	?	?	?	?
Choudhuri 2008	?	?	?	?	?	?
Chu 2006	?	?	?	?	?	?
Cocelli 2012	?	?	?	?	?	?
Demiraran 2005	?	?	?	?	?	?
Demiraran 2006	?	?	?	?	?	?
Doda 2009	?	?	?	?	?	?
Ekemen 2008	?	?	?	?	?	?
El Deeb 2011	?	?	?	?	?	?
Engelhardt 2003	?	?	?	?	?	?
Ertugrul 2006	?	?	?	?	?	?
Fan 2000	?	?	?	?	?	?
Friedrichsdorf 2016	?	?	?	?	?	?
Giraldes 2016	?	?	?	?	?	?
Gunduz 2001	?	?	?	?	?	?
Gunduz 2006	?	?	?	?	?	?
Güneş 2004	?	?	?	?	?	?
Gupta 2003	?	?	?	?	?	?
Gupta 2009	?	?	?	?	?	?
Gupta 2017	?	?	?	?	?	?
Hegazy 2013	?	?	?	?	?	?
Honarmand 2013	?	?	?	?	?	?
Honarmand 2015	?	?	?	?	?	?
Hullett 2006	?	?	?	?	?	?
Inanoglu 2010	?	?	?	?	?	?
Khan 2008	?	?	?	?	?	?
Khan 2015	?	?	?	?	?	?
Khosravi 2006	?	?	?	?	?	?
Krishnadas 2016	?	?	?	?	?	?
Lalq 2009	?	?	?	?	?	?
Liaquat 2017	?	?	?	?	?	?
Moyado Garcia 2009	?	?	?	?	?	?
Numanoğlu 2014	?	?	?	?	?	?
Ozalevli 2005	?	?	?	?	?	?
Ozcengiz 2001	?	?	?	?	?	?
Ozer 2003	?	?	?	?	?	?
Ozkan 2003	?	?	?	?	?	?
Ozköse 2000	?	?	?	?	?	?
Pendeville 2000	?	?	?	?	?	?
Prakash 2006	?	?	?	?	?	?
Senel 2001	?	?	?	?	?	?
Shrestha 2010	?	?	?	?	?	?
Solanki 2016	?	?	?	?	?	?
Taheri 2010	?	?	?	?	?	?
Tekelioğlu 2013	?	?	?	?	?	?
Ugur 2008	?	?	?	?	?	?
Ugur 2013	?	?	?	?	?	?
Umuroglu 2004	?	?	?	?	?	?
Uysal 2011	?	?	?	?	?	?
Viltanen 2001	?	?	?	?	?	?
Yenigun 2015	?	?	?	?	?	?
Yildiz 2010	?	?	?	?	?	?

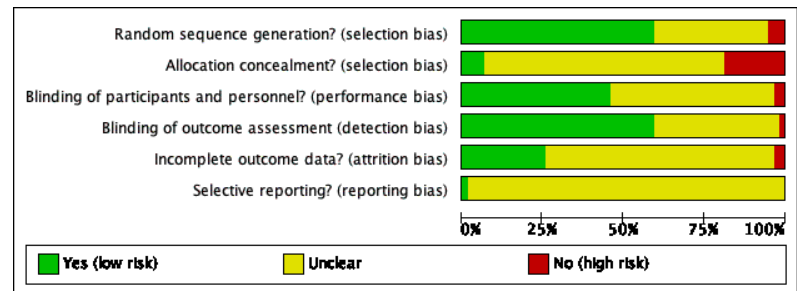


Figure 4. Risk of bias of included studies

Review authors' judgement of risk of bias for all included studies. Detail of each risk of bias category for each study and as percentages across all studies.

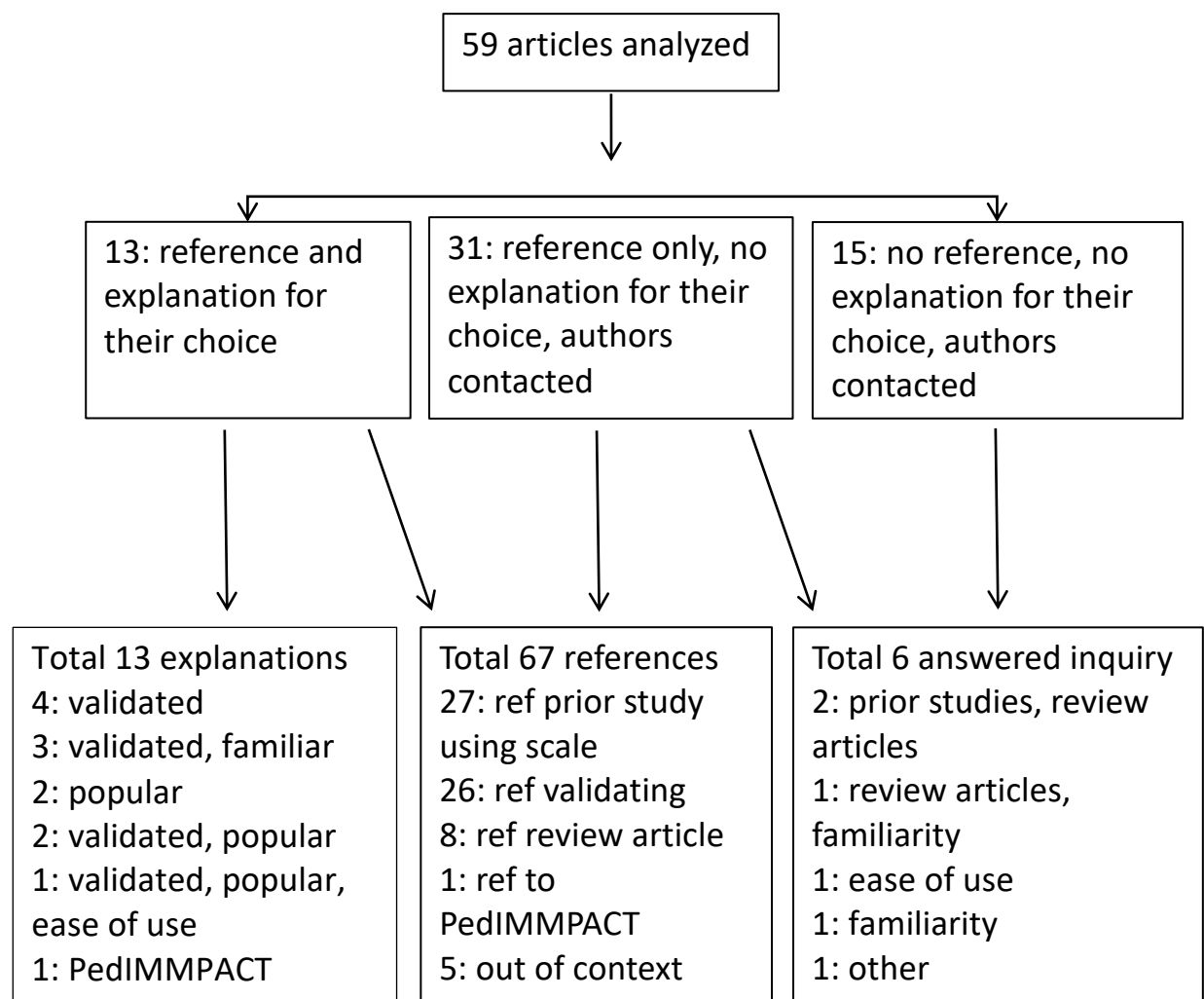


Figure 5. Summary of choice of pain score used in trials

Explanation of choice of pain scales used, pain scale references and author response to enquiry regarding choice.

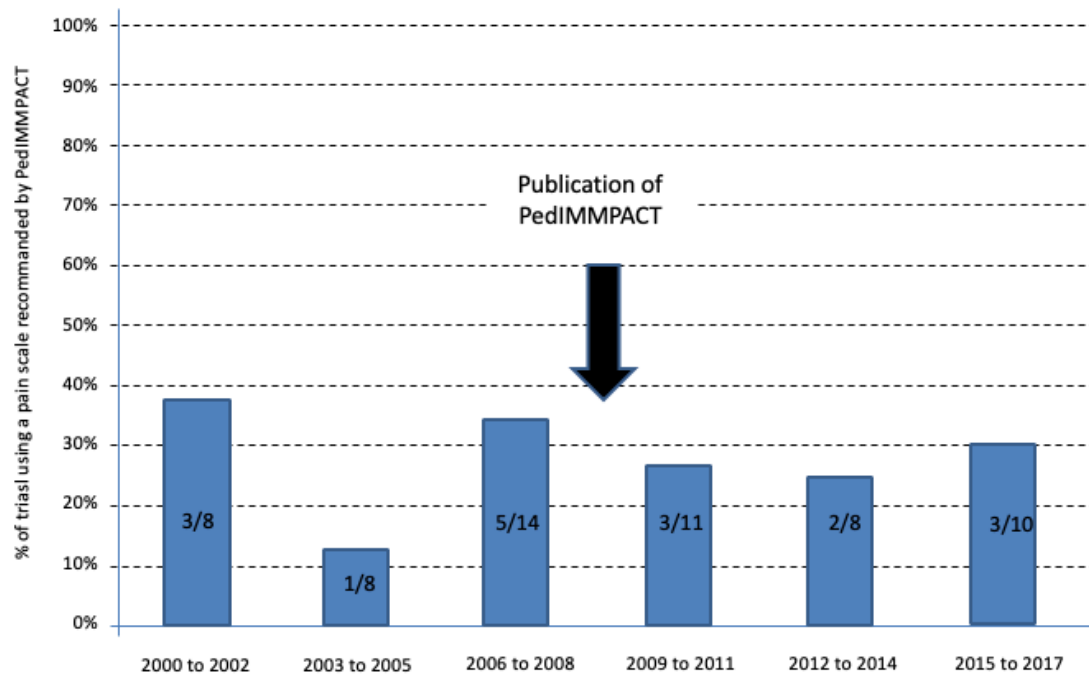


Figure 6. Summary of the proportion of trials complying with PedIMMPACT, per 3-year period

13 Tableaux

Table 1. Summary of pain scales recommended by PedIMMPACT

Table 2. Scoring system for the Face, Legs, Activity, Cry, Consolability pain scale

Table 3. Scoring system for the Children's Hospital of Eastern Ontario Pain Scale

Table 4. Scoring criteria for the Toddler Preschooler Postoperative Pain Scale.

Table 5. Questions for the Parents' Postoperative Pain Measure

Table 6. Summary of baseline data and pain scores used

Table 7. Authors response to inquiry regarding their choice of pain score

	Age	Measure	Context
Self-reporting scales	3 to 4 years	Poker Chip Tool	Procedure and post op
	4 to 12 years	Faces Pain Scale Revised	Procedure and post op
	≥8 years	Visual Analog Scale	Procedure and post op
Observational scales	≥1 year	FLACC	Procedure and post op
	≥1 year	CHEOPS	Procedure and post op
	1 to 5 years	TPPPS	Post op
	≥1 year	PPPM	Post op at home

Table 1. Summary of pain scales recommended by PedIMMPACT

Modified from (23).

Key FLACC: face, legs, activity, cry, consolability, CHEOPS: Children's Hospital of Eastern Ontario Pain Scale pain score, TPPPS: Toddler-Preschool Postoperative Pain Scale, PPPM: Parents' Postoperative Pain Measure

	Score		
Criteria	0	1	2
Face	No particular expression or smile	Occasional grimace or frown, withdrawn, uninterested	Frequent to constant quivering chin, clenched jaw
Legs	Normal position or relaxed	Uneasy, restless, tense	Kicking or legs drawn up
Activity	Lying quietly, normal position, moves easily	Squirming, shifting, back and forth, tense	Arched, rigid or jerking
Cry	No cry (awake or asleep)	Moans or whimpers, occasional complaint	Crying steadily, screams or sobs, frequent complaints
Consolability	Content, relaxed	Reassured by occasional touching, hugging or being talked to, distractible	Difficult to console or comfort

Table 2. Scoring system for the Face, Legs, Activity, Cry, Consolability pain scale

Modified from (31)

Behavior	Score	Criteria
Cry	1	No cry
	2	Moaning or crying
	3	Screaming
Facial	0	Smiling
	1	Composed
	2	Grimace
Child Verbal	0	Positive
	1	No complaints or other complaints
	2	Pain complaints
Torso	1	Neutral
	2	Shifting, tense, shivering, upright or restrained
Touch	1	Not touching
	2	Reaching, touching, grabbing or restrained
Legs	1	Neutral
	2	Squirming/kicking, drawn up/tensed, standing, restrained

Table 3. Scoring system for the Children's Hospital of Eastern Ontario Pain Scale

Modified from (5)

	Expression	Yes (1)/No (0)
Verbal Pain Expression	Verbal pain complaint or cry	
	Groan moan or grunt	
	Scream	
Facial Pain Expression	Open mouth, lips pulled back at corners	
	Squint or closed eyes	
	Brow bulge or forehead furrow	
Bodily Pain Expression	Restless motor behavior - Rub or touch painful area	

Table 4. Scoring criteria for the Toddler Preschooler Postoperative Pain Scale.

Modified from (32)

Behavior	Yes (1)/No (0)
Whine or complains more than usual?	
Cries more easily than usual?	
Plays less than usual?	
Does not do the things s/he normally does?	
Acts more worried than usual?	
Acts more quiet than usual?	
Has less energy than usual?	
Refuses to eat?	
Eats less than usual?	
Holds the sore part of his/her body?	
Tries not to bump the sore part of his/her body?	
Groans or moans more than usual?	
Looks more flushed than usual?	
Wants to be close to you more than usual?	
Takes medication when s/he normally refuses?	

Table 5. Questions for the Parents' Postoperative Pain Measure
Modified from (33)

First Author	Year	Country	Age (years unless specified)	Surgery Type	Number of Pain Scales Used	Pain Intensity Scale Name	Scale	Complies with PedIMMPACT
Bedirli	2017	Turkey	2 to 12	ENT	1	Pain Discomfort Scale	n/a	No
Gupta	2017	India	1 to 8	Infraumbilical	1	FLACC	0-10	Yes
Liaqat	2017	Pakistan	1 to 12	Infraumbilical	1	Wong-Baker Faces	0-5	No
Giraldes	2016	Brazil	3 to 12	Infraumbilical	2	Wong-Baker Faces	0-5	No
						Simple Descriptive Scale	0-3	No
Krishnadas	2016	India	1 to 5	Infraumbilical	1	Modified CHEOPS	0-10	No
Solanki	2016	India	1 to 12	Infraumbilical	1	FLACC	0-10	Yes
Friedrichsdorf	2015	USA	4 to 15	ENT	2	Faces Pain Scale Revised	0-10	Yes
						NRS-11	0-10	No
Honarmand	2015	Iran	5 to 15	ENT	1	Objective Pain Scale	0-10	No
Khan	2015	Pakistan	3 to 8	Infraumbilical	1	Modified Objective Pain Scale	0-10	No
Yenigun	2015	Turkey	5 to 15	ENT	1	Modified CHEOPS	0-10	No
Numanoğlu	2014	Turkey	2 to 7	Infraumbilical	2	Faces Pain Scale Revised	0-10	Yes
						Pain Discomfort Scale	0-10	No
Aydin	2013	Turkey	3 to 16	ENT	1	Verbal Rating Scale	0-10	No
Hegazy	2013	Egypt	1 to 4	Abdominal	1	FLACC	0-10	Yes
Honarmand	2013	Iran	2 to 15	ENT	1	Modified CHEOPS	0-10	No
Tekelioglu	2013	Turkey	4 to 10	ENT	1	Wong-Baker Faces	0-5	No
Ugur	2013	Turkey	3 to 10	ENT	1	Modified CHEOPS	0-10	No
Ayatollahi	2012	Iran	5 to 12	ENT	1	Modified CHEOPS	0-10	No
Cocelli	2012	Turkey	2 to 9	ENT	1	Maunuksela pain score	0-10	No
El Deeb	2011	Egypt	2 to 12	Infraumbilical	1	FLACC	0-10	Yes
Uysal	2011	Turkey	6 to 16	ENT	1	Pain Discomfort Scale	0-10	No
Akbay	2010	Turkey	4 to 15	ENT	1	McGrath's Face Scale	0.04-0.97	No

Inanoglu	2010	Turkey	2 to 12	Abdominal	1	Modified CHEOPS	0-10	No
Shrestha	2010	Nepal	1 to 6	Infraumbilical	1	Modified Pain Discomfort Scale	1-6	No
Taheri	2010	Iran	1 to 7	Infraumbilical	1	FLACC	0-10	Yes
Yildiz	2010	Turkey	1 to 7	Infraumbilical	1	CHIPPS	0-10	No
Doda	2009	India	2 to 5	Infraumbilical	1	Modified TPPPS	0-10	No
Gupta	2009	India	1 to 12	Infraumbilical	1	Objective Pain Score	n/a	No
Laiq	2009	Pakistan	2 to 6	Infraumbilical	1	Wong-Baker Faces	0-10	No
Moyao-García	2009	Mexico	1 to 10	Various	3	CHEOPS	4-13	Yes
						VAS	0-100	No
						Faces Pain Scale	0-5	No
Ali	2008	Iran	4 to 10	ENT	1	Faces Pain Scale Revised	0-10	Yes
Atef	2008	Egypt	3 to 10	ENT	1	Objective Pain Scale	0-10	No
Choudhuri	2008	India	3 to 9	Infraumbilical	1	AIIMS	0-10	No
Ekemen	2008	Turkey	2 to 12	Infraumbilical	1	Four-point restlessness–pain protocol	0-3	No
Khan	2008	Pakistan	13 to 53 months	Infraumbilical	1	Modified TPPPS	0-10	No
Ugur	2008	Turkey	7 to 16	ENT	1	VAS	0-10	Yes
Antila	2006	Finland	9 to 15	ENT	1	VAS	0-10	Yes
Chu	2006	Taiwan	1 to 6	cardiac	1	CHEOPS	4-13	Yes
Demiraran	2006	Turkey	1 to 6	Infraumbilical	1	Faces Pain Scale	0-5	No
Ertugrul	2006	Turkey	1 to 7	ENT	1	Modified TPPPS	0-10	No
Gunduz	2006	Turkey	1 to 10	Infraumbilical	1	Modified CHEOPS	0-10	No
Hullett	2006	Australia	1 to 8	ENT	1	Prince Margaret Hospital pain score	0-10	No
Khosravi	2006	Iran	2 to 7	Infraumbilical	2	CHEOPS	4-13	Yes
						Categorical Pain Scale (by parents)	n/a	No
Prakash	2006	India	2 to 8	Infraumbilical	1	AIIMS	0-10	No

Demiraran	2005	Turkey	7 to 14	Infraumbilical	1	Modified Objective Pain Scale	0-10	No
Ozalevli	2005	Turkey	6 to 12	ENT	1	Modified CHEOPS	0-10	No
Güneş	2004	Turkey	1 to 10	Infraumbilical	1	CHEOPS	4-13	Yes
Umuroglu	2004	Turkey	5 to 12	ENT	2	NRS	0-5	No
						CHEOPS	n/a	Unclear
Engelhardt	2003	Scotland	2 to 14	ENT	1	Pain Score	1-5	No
Gupta	2003	India	1 to 12	Infraumbilical	1	Objective Pain Scale	0-10	No
Ozer	2003	Turkey	4 to 7	ENT	1	Faces Pain Scale	0-6	No
Ozkan	2003	Turkey	1 to 5	Infraumbilical	1	Modified TPPPS	0-10	No
Gunduz	2001	Turkey	1 to 5	Infraumbilical	1	Broadman Objective Pain Discomfort scale	0-10	No
Ozcengiz	2001	Turkey	4 to 10	Infraumbilical	1	Objective Pain Scale	0-10	No
Senel	2001	Turkey	12–84 months	Infraumbilical	1	Three-point Pain and Demeanour Scale	0-3	No
Viitanen	2001	Finland	1 to 3	ENT	2	Three objective components of Hannallah Pain Discomfort Scale	0-6	No
						home post-operative pain questionnaire (parents)	n/a	Unclear
Chiaretti	2000	Italy	4–196 months	Neurosurgery	3	CHEOPS	4-13	Yes
						McGrath Facial Scale	0-8	No
						Haemodynamic score	n/a	No
Fan	2000	Taiwan	1 to 8	Infraumbilical	2	CHEOPS used in Taiwan	0-4	No
						Verbal expression	0-2	No
Ozköse	2000	Turkey	2 to 11	ENT	2	VAS	0-10	Yes
						Faces Pain Scale	0-6	No
Pendeville	2000	Belgium	2 to 9	ENT	3	CHEOPS	4-13	Yes
						Pain on swallowing	0-3	No
						5 - point scale	1-5	No

Table 6. Summary of baseline data and pain scores used

Key: CHEOPS (Children's Hospital of Eastern Ontario Pain Scale) pain score, FLACC (face, legs, activity, cry, consolability), modified TPPPS (Toddler-Preschool Postoperative Pain Scale) AIIMS (All India Institute of Medical Sciences) pain discomfort scale, the CHIPPS scale (Children's and Infants' Postoperative Pain Scale), VAS (visual analogue scale), NRS (numerical rating scale); n/a = unspecified; ENT: (Ear, Nose and Throat)

Response from authors regarding choice of pain score
What we prefer is the easy accessibility of the scale and easy implementation which have less complicated variables.
We used numeric pain score corresponding to the visual pain scales such as CHEOPS, or Wong-Baker's scales in patients not able to express their pain. But we asked if patient was cooperative and able to express their pain in VAS scale as performed in adult patients. We currently used Vas or numeric pain scale (especially elder patients who are not able to express their pain) for adult and elderly. We commonly used pain scales based on facial expression as indicated above in pediatric patients, but contrary to my previous affiliation, our pediatric patients are low due to separate hospitals.
Especially prior published articles and the guidelines which had cited.
We have chosen the CHEOPS with the aid of prior articles, and systematic reviews.
To my knowledge, the factor aided us in choosing the pain score used in our study was the current pain score used in hospital based on the reference anesthesiology books.
I published the paper a long time back in India and currently I work in UK but not in anesthesia anymore, so I don't have the details of that paper.

Table 7. Authors response to inquiry regarding their choice of pain score

14 Annexe: Références bibliographiques des articles inclus dans l'analyse transversale

En ordre alphabétique du premier auteur:

Akbay BK, Yildizbas S, Guclu E, Yilmaz S, Iskender A, Ozturk O. Analgesic efficacy of topical tramadol in the control of postoperative pain in children after tonsillectomy. *J Anesth.* 2010;24(5):705-8.

Ali SM, Shahrabano S, Ulhaq TS. Tramadol for pain relief in children undergoing adenotonsillectomy: a comparison with dextromethorphan. *Laryngoscope.* 2008;118(9):1547-9.

Antila H, Manner T, Kuurila K, Salanterä S, Kujala R, Aantaa R. Ketoprofen and tramadol for analgesia during early recovery after tonsillectomy in children. *Paediatr Anaesth.* 2006;16(5):548-53.

Atef A, Fawaz AA. Peritonsillar infiltration with tramadol improves pediatric tonsillectomy pain. *Eur Arch Otorhinolaryngol.* 2008;265(5):571-4.

Ayatollahi V, Behdad S, Hatami M, Moshtaghiun H, Baghianimoghadam B. Comparison of peritonsillar infiltration effects of ketamine and tramadol on post tonsillectomy pain: a double-blinded randomized placebo-controlled clinical trial. *Croat Med J.* 2012;53(2):155-61.

Aydin G, Apan A, Kose E, Oz G. Pre-versus postoperative tramadol instillation: both are effective for decreasing pain and/or agitation in pediatric adenotonsillectomy. *Turkish Journal of Medical Sciences.* 2013;43 899-904.

Bedirli N, Akçabay M, Emik U. Tramadol vs dexmedetomidine for emergence agitation control in pediatric patients undergoing adenotonsillectomy with sevoflurane anesthesia: prospective randomized controlled clinical study. *BMC Anesthesiol.* 2017;17(1):41.

Chiaretti A, Viola L, Pietrini D, Piastra M, Savioli A, Tortorolo L, et al. Preemptive analgesia with tramadol and fentanyl in pediatric neurosurgery. *Childs Nerv Syst.* 2000;16(2):93-9; discussion 100.

Choudhuri AH, Dharmani P, Kumar N, Prakash A. Comparison of caudal epidural bupivacaine with bupivacaine plus tramadol and bupivacaine plus ketamine for postoperative analgesia in children. *Anaesth Intensive Care.* 2008;36(2):174-9.

Chu YC, Lin SM, Hsieh YC, Chan KH, Tsou MY. Intraoperative administration of tramadol for postoperative nurse-controlled analgesia resulted in earlier awakening and less sedation than morphine in children after cardiac surgery. *Anesth Analg.* 2006;102(6):1668-73.

Cocelli LP, Ugur BK, Durucu C, Kul S, Arik H, Mumbuc S. Comparison of pre-emptive tonsillar lodge infiltration with ropivacaine versus intravenous tramadol in pediatric tonsillectomies: a randomized placebo-controlled study. *Int J Pediatr Otorhinolaryngol.* 2012;76(5):653-7.

Demiraran Y, Kocaman B, Akman RY. A comparison of the postoperative analgesic efficacy of single-dose epidural tramadol versus morphine in children. *Br J Anaesth.* 2005;95(4):510-3.

Demiraran Y, Ilce Z, Kocaman B, Bozkurt P. Does tramadol wound infiltration offer an advantage over bupivacaine for postoperative analgesia in children following herniotomy? *Paediatr Anaesth.* 2006;16(10):1047-50.

Doda M, Mukherjee S. Postoperative Analgesia in Children- Comparative Study between Caudal Bupivacaine and Bupivacaine plus Tramadol. *Indian J Anaesth.* 2009;53(4):463-6.

Ekemen S, Yelken B, Ilhan H, Tokar B. A comparison of analgesic efficacy of tramadol and pethidine for management of postoperative pain in children: a randomized, controlled study. *Pediatr Surg Int.* 2008;24(6):695-8.

El Deeb A, El-Morsy GZ. Comparison of preemptive analgesic effect of intravenous ketorolac versus tramadol in pediatric inguinal herniotomy: A randomized double blind study. *Egyptian Journal of Anaesthesia.* 2011;27(4):207-11.

- Engelhardt T, Steel E, Johnston G, Veitch DY. Tramadol for pain relief in children undergoing tonsillectomy: a comparison with morphine. *Paediatr Anaesth*. 2003;13(3):249-52.
- Ertugrul F, Akbas M, Karsli B, Kayacan N, Bulut F, Trakya A. Pain relief for children after adenotonsillectomy. *J Int Med Res*. 2006;34(6):648-54.
- Fan KT, Lee TH, Yu KL, Tang CS, Lu DV, Chen PY, et al. Influences of tramadol on emergence characteristics from sevoflurane anesthesia in pediatric ambulatory surgery. *Kaohsiung J Med Sci*. 2000;16(5):255-60.
- Friedrichsdorf SJ, Postier AC, Foster LP, Lander TA, Tibesar RJ, Lu Y, et al. Tramadol versus codeine/acetaminophen after pediatric tonsillectomy: A prospective, double-blinded, randomized controlled trial. *J Opioid Manag*. 2015;11(4):283-94.
- Giraldes AL, Sousa AM, Slullitel A, Guimarães GM, Santos MG, Pinto RE, et al. Tramadol wound infiltration is not different from intravenous tramadol in children: a randomized controlled trial. *J Clin Anesth*. 2016;28:62-6.
- Gunduz M, Ozalevli M, Ozbek H, Ozcengiz D. Comparison of caudal ketamine with lidocaine or tramadol administration for postoperative analgesia of hypospadias surgery in children. *Paediatr Anaesth*. 2006;16(2):158-63.
- Gunduz M, Ozcengiz D, Ozbek H, Isik G. A comparison of single dose caudal tramadol, tramadol plus bupivacaine and bupivacaine administration for postoperative analgesia in children. *Paediatr Anaesth*. 2001;11(3):323-6.
- Güneş Y, Seçen M, Ozcengiz D, Gündüz M, Balcioglu O, Işık G. Comparison of caudal ropivacaine, ropivacaine plus ketamine and ropivacaine plus tramadol administration for postoperative analgesia in children. *Paediatr Anaesth*. 2004;14(7):557-63.
- Gupta S, Sharma R. Comparison of analgesic efficacy of caudal dexmedetomidine versus caudal tramadol with ropivacaine in paediatric infraumbilical surgeries: A prospective, randomised, double-blinded clinical study. *Indian J Anaesth*. 2017;61(6):499-504.

Gupta R, Kaur G, Aujla K, Kumari A. Comparison of efficacy of caudal bupivacaine alone or in combination with butorphanol or tramadol for postoperative analgesia in children undergoing infra-umbilical surgeries. *J Anaesth Clin Pharmacol*. 2009;25(2):187-9.

Gupta R, Jain A, Bhardwaj M, Rattan K, Chhabra B, Gupta M. Postoperative pain relief in children with caudal block using bupivacaine, tramadol & combination of both- A comparison. *J Anaesth Clin Pharmacol*. 2003;19(4):383-7.

Hegazy M, Ghoneim AA. Comparative study of caudal bupivacaine versus bupivacaine with tramadol for postoperative analgesia in paediatric cancer patients. *Chin -Ger J Clin Oncol*. 2013;12(3):118–22.

Honarmand A, Safavi MR, Jamshidi M. The preventative analgesic effect of preincisional peritonsillar infiltration of two low doses of ketamine for postoperative pain relief in children following adenotonsillectomy. A randomized, double-blind, placebo-controlled study. *Paediatr Anaesth*. 2008;18(6):508-14.

Honarmand A, Safavi M, Kashefi P, Hosseini B, Badiei S. Comparison of effect of intravenous ketamine, peritonsillar infiltration of tramadol and their combination on pediatric posttonsillectomy pain: A double-blinded randomized placebo-controlled clinical trial. *Res Pharm Sci*. 2013;8(3):177-83.

Hullett BJ, Chambers NA, Pascoe EM, Johnson C. Tramadol vs morphine during adenotonsillectomy for obstructive sleep apnea in children. *Paediatr Anaesth*. 2006;16(6):648-53.

Inanoglu K, Ozcengiz D, Gunes Y, Unlugenc H, Isik G. Epidural ropivacaine versus ropivacaine plus tramadol in postoperative analgesia in children undergoing major abdominal surgery: a comparison. *J Anesth*. 2010;24(5):700-4.

Khan S, Memon MI. Comparison of caudal bupivacaine and bupivacaine-tramadol for postoperative analgesia in children with hypospadias repair. *J Coll Physicians Surg Pak*. 2008;18(10):601-4.

Khan Q, Nasir S, Ahmed B, Shehbaz L, Basar S. To Compare the Duration and side effects of Caudal Tramadol Bupivacaine versus Ketamine – Bupivacaine for

Postoperative analgesia in the pediatric age group. *Pakistan Journal of Medical and Health Sciences*. 2015;9(4):1233- 6.

Khosravi MB, Khezri S, Azemati S. Tramadol for pain relief in children undergoing herniotomy: a comparison with ilioinguinal and iliohypogastric blocks. *Paediatr Anaesth*. 2006;16(1):54-8.

Krishnadas A, Suvarna K, Hema VR, Taznim M. A comparison of ropivacaine, ropivacaine with tramadol and ropivacaine with midazolam for post-operative caudal epidural analgesia. *Indian J Anaesth*. 2016;60(11):827-32.

Laiq N, Khan MN, Tahmeedullah, Gandapur YK, Khan S. Comparison of caudal bupivacaine and bupivacaine-tramadol for postoperative analgesia in children undergoing hypospadias surgery. *J Coll Physicians Surg Pak*. 2009;19(11):678-81.

Liaqat N, Dar SH. Comparison of single-dose nalbuphine versus tramadol for postoperative pain management in children: a randomized, controlled trial. *Korean J Anesthesiol*. 2017;70(2):184-7.

Moyao-García D, Hernández-Palacios JC, Ramírez-Mora JC, Nava-Ocampo AA. A pilot study of nalbuphine versus tramadol administered through continuous intravenous infusion for postoperative pain control in children. *Acta Biomed*. 2009;80(2):124-30.

Numanoğlu KV, Ayoğlu H, Er DT. Efficacy of tramadol as a preincisional infiltration anesthetic in children undergoing inguinal hernia repair: a prospective randomized study. *Ther Clin Risk Manag*. 2014;10:753-8.

Ozalevli M, Unlügenç H, Tuncer U, Güneş Y, Ozcengiz D. Comparison of morphine and tramadol by patient-controlled analgesia for postoperative analgesia after tonsillectomy in children. *Paediatr Anaesth*. 2005;15(11):979-84.

Ozcengiz D, Gunduz M, Ozbek H, Isik G. Comparison of caudal morphine and tramadol for postoperative pain control in children undergoing inguinal herniorrhaphy. *Paediatr Anaesth*. 2001;11(4):459-64.

- Ozer Z, Görür K, Altunkan AA, Bilgin E, Camdeviren H, Oral U. Efficacy of tramadol versus meperidine for pain relief and safe recovery after adenotonsillectomy. *Eur J Anaesthesiol.* 2003;20(11):920-4.
- Ozkan S, Poçan S, Bahar A, Cakir O, Gökben M. The effect of caudal bupivacaine versus tramadol in post-operative analgesia for paediatric patients. *J Int Med Res.* 2003;31(6):497-502.
- Ozköse Z, Akçabay M, Kemaloğlu YK, Sezenler S. Relief of posttonsillectomy pain with low-dose tramadol given at induction of anesthesia in children. *Int J Pediatr Otorhinolaryngol.* 2000;53(3):207-14.
- Pendeville PE, Von Montigny S, Dort JP, Veyckemans F. Double-blind randomized study of tramadol vs. paracetamol in analgesia after day-case tonsillectomy in children. *Eur J Anaesthesiol.* 2000;17(9):576-82.
- Prakash S, Tyagi R, Gogia AR, Singh R. Efficacy of three doses of tramadol with bupivacaine for caudal analgesia in paediatric inguinal herniotomy. *Br J Anaesth.* 2006;97(3):385-8.
- Senel AC, Akyol A, Dohman D, Solak M. Caudal bupivacaine-tramadol combination for postoperative analgesia in pediatric herniorrhaphy. *Acta Anaesthesiol Scand.* 2001;45(6):786-9.
- Shrestha SK, Bhattarai B. Caudal bupivacaine vs bupivacaine plus tramadol in post-operative analgesia in children. *J Nepal Health Res Counc.* 2010;8(2):99-102.
- Solanki NM, Engineer SR, Jansari DB, Patel RJ. Comparison of caudal tramadol versus caudal fentanyl with bupivacaine for prolongation of postoperative analgesia in pediatric patients. *Saudi J Anaesth.* 2016;10(2):154-60.
- Taheri R, Shayeghi S, Razavi SS, Sadeghi A, Ghabili K, Ghojzadeh M, et al. Efficacy of bupivacaine-neostigmine and bupivacaine-tramadol in caudal block in pediatric inguinal herniorrhaphy. *Paediatr Anaesth.* 2010;20(9):866-72.
- Tekelioglu UY, Apuhan T, Akkaya A, Demirhan A, Yildiz I, Simsek T, et al. Comparison of topical tramadol and ketamine in pain treatment after tonsillectomy. *Paediatr Anaesth.* 2013;23(6):496-501.

- Ugur KS, Karabayirli S, Demircioğlu R, Ark N, Kurtaran H, Muslu B, et al. The comparison of preincisional peritonsillar infiltration of ketamine and tramadol for postoperative pain relief on children following adenotonsillectomy. *Int J Pediatr Otorhinolaryngol.* 2013;77(11):1825-9.
- Ugur MB, Yilmaz M, Altunkaya H, Cinar F, Ozer Y, Beder L. Effects of intramuscular and peritonsillar injection of tramadol before tonsillectomy: a double blind, randomized, placebo-controlled clinical trial. *Int J Pediatr Otorhinolaryngol.* 2008;72(2):241-8.
- Umuroğlu T, Eti Z, Ciftçi H, Yilmaz Göğüş F. Analgesia for adenotonsillectomy in children: a comparison of morphine, ketamine and tramadol. *Paediatr Anaesth.* 2004;14(7):568-73.
- Uysal HY, Takmaz SA, Yaman F, Baltacı B, Başar H. The efficacy of intravenous paracetamol versus tramadol for postoperative analgesia after adenotonsillectomy in children. *J Clin Anesth.* 2011;23(1):53-7.
- Viitanen H, Annala P. Analgesic efficacy of tramadol 2 mg kg⁻¹ for paediatric day-case adenoidectomy. *Br J Anaesth.* 2001;86(4):572-5.
- Yenigun A, Et T, Aytac S, Olcay B. Comparison of different administration of ketamine and intravenous tramadol hydrochloride for postoperative pain relief and sedation after pediatric tonsillectomy. *J Craniofac Surg.* 2015;26(1):e21-4.
- Yildiz TS, Ozdamar D, Bagus F, Solak M, Toker K. Levobupivacaine-tramadol combination for caudal block in children: a randomized, double-blinded, prospective study. *Paediatr Anaesth.* 2010;20(6):524-9.