Received: 10 February 2022 | Accepted: 10 April 2022

DOI: 10.1111/pde.15015

ORIGINAL ARTICLE

Pediatric Dermatology WILEY

Inter- and intra-observer variability in the selection of therapy for infantile hemangiomas among pediatric dermatologists in Spain

María Colmenero MD[1] | Javier del Boz MD, PhD[2] | José Bernabeu Wittel MD, PhD[3] | Esther Roé MD, PhD[4] | Marta Feito-Rodríguez MD, PhD[5] | María Asunción Vicente-Villa MD[6] | Ana Martín-Santiago MD[7] | Sara Isabel Palencia Pérez MD, PhD[8] | Antoni Azon MD, PhD[9] | Marta Valdivielso-Ramos MD, PhD[10] | Antonio Torrelo MD, PhD[11] | Ana Isabel Sánchez Moya MD, PhD[12] | Minia Campos-Domínguez MD, PhD[13] | Gloria Garnacho-Saucedo MD[14] | José Manuel Azaña Defez MD, PhD[15] | Ángel Vera Casaño MD[2] | Jesús Tercedor-Sánchez MD[16] | Rebeca Alcalá MD[17] | María Antonia González-Enseyat MD[6] | Aniza Giacaman MD[7] | Ángela Hernández-Martin MD, PhD[11] | María Teresa Monserrat García MD[3] | Ana Bauzá MD[7] | Javier Domínguez-Cruz MD, PhD[3] | Ignacio García-Doval MD, MSc, PhD[18] | Mercè Grau-Pérez MD, MSc[18]

[1]Dermatology Department, Hospital Costa del Sol, Marbella, Spain

[2]Dermatology Department, Hospital Universitario Regional, Málaga, Spain

[3]Dermatology Department, Hospital Universitario Virgen del Rocío, Sevilla, Spain

[4]Dermatology Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

[5]Dermatology Department, Hospital Universitario La Paz, Madrid, Spain

[6]Dermatology Department, Hospital Sant Joan de Deu, Barcelona, Spain

[7]Dermatology Department, Hospital Son Espases, Mallorca, Spain

[8]Dermatology Department, Hospital Universitario 12 de Octubre, Madrid, Spain

[9]Dermatology Department, Hospital Universitari Sant Joan de Reus, Reus, Spain

[10]Dermatology Department, Hospital Universitario Infanta Leonor, Madrid, Spain

[11]Dermatology Department, Hospital Infantil Universitario Niño Jesús, Madrid, Spain

[12]Dermatology Department, Complejo Hospitalario Universitario de Toledo, Toledo, Spain

[13]Dermatology Department, Hospital Universitario Gregorio Marañón, Madrid, Spain

[14]Dermatology Department, Hospital Universitario Reina Sofia, Córdoba, Spain

[15]Dermatology Department, Complejo Hospitalario Universitario de Albacete, Albacete, Spain

[16]Dermatology Department, Hospital Universitario Virgen de las Nieves, Granada, Spain

[17]Dermatology Department, Hospital de Sagunto, Valencia, Spain

[18]Research Unit, Fundación Piel Sana AEDV, Madrid, Spain

Correspondence
María Colmenero, Dermatology Department, Hospital Costa del Sol, Marbella, Spain.
Email: sendracolmenero@gmail.com

Funding information
This study was funded by the Fundación Piel Sana AEDV (Healthy Skin Foundation of the Spanish Academy of Dermatology and Venereology).

Abstract

Background: Guidelines and expert recommendations on infantile hemangiomas (IH) are aimed at increasing homogeneity in clinical decisions based on the risk of sequelae.

Objective: The objective was to analyze the inter- and intra-observer agreement among pediatric dermatologists in the choice of treatment for IH.

Methods: We performed a cross-sectional inter-rater and intra-rater agreement study within the Spanish infantile hemangioma registry. Twenty-seven pediatric dermatologists were invited to participate in a survey with 50 clinical vignettes randomly selected within the registry. Each vignette contained a picture of an infantile hemangioma with a clinical description. Raters chose therapy among observation, topical timolol, or oral propranolol. The same survey reordered was completed 1 month later to assess intra-rater agreement. Vignettes were stratified into hemangioma risk categories following the Spanish consensus on IH. The agreement was measured using kappa statistics appropriate for the type of data (Gwet's AC_1 coefficient and Gwet's paired *t* test).

Results: Twenty-four dermatologists completed the survey. Vignettes represented 7.8% of the Spanish hemangioma registry. The inter-rater agreement on the treatment decision was fair ($AC_1 = 0.39$, 95% confidence interval [CI]: 0.30–0.47). When stratified by risk category, good agreement was reached for high-risk hemangiomas ($AC_1 = 0.77$, 95% CI: 0.51–1.00), whereas for intermediate- and low-risk categories, the agreement was only fair (AC_1 0.31, 95% CI: 0.16–0.46 and $AC_1 = 0.38$, 95% CI: 0.27–0.48, respectively). Propranolol was the main option for high-risk hemangiomas (86.4%), timolol for intermediate-risk (36.8%), and observation for low-risk ones (55.9%). The intra-rater agreement was good. The inter-rater agreement between pediatric dermatologists on the treatment of IH is only fair. Variability was most significant with intermediate- and low-risk hemangiomas.

KEYWORDS
cross-sectional studies, dermatology, hemangioma, observer variation, pediatrics, propranolol, surveys and questionnaires, timolol

1 | INTRODUCTION

Infantile hemangiomas (IH) are the most common benign vascular tumor in childhood with a prevalence of 2.6%–4.5% in newborns.[1] Beta-blockers have become the standard treatment for IH. Randomized clinical trials support the use of oral propranolol, while clinical trials supporting the use of topical timolol are scarce and focus on superficial hemangiomas in the early proliferative stage.[2] Currently, several clinical practice guidelines offer recommendations for the management of IH, based on the propensity for the development of complications.[3–5] One of the aims of clinical practice guidelines is to decrease variation in the care provided in similar clinical scenarios.

The objective of this study was to analyze the agreement among pediatric dermatologists in the treatment of IH, which has not been studied previously.

2 | MATERIALS AND METHODS

The Spanish Academy of Dermatology and Venereology (AEDV) started an IH nationwide prospective cohort in 2016 that recruited all consecutive pediatric patients diagnosed with IH in 12 Spanish hospitals between June 2016 and October 2019 ($N = 640$). The study was approved by the Research Ethics Committee of the Hospital Santa Creu i Sant Pau (Barcelona, 16/079). The patients' representatives gave written informed

consent for the use of clinical images for research purposes. The present study was conducted from this cohort between March and July 2021 following the Guidelines for Reporting Reliability and Agreement Studies.[6]

2.1 | Study design

This is a cross-sectional inter-rater and intra-rater agreement study on the therapeutic treatment of IH among pediatric dermatologists in Spain.

2.2 | Rater population of interest and subjects

All pediatric dermatologists working in Spain and participating as researchers in the nationwide IH cohort were invited to participate in the study. Infants younger than 9 months of age with IH referred for specialized dermatology care were included in the AEDV IH cohort. To obtain a representative sample of the cohort, 50 cases were selected by simple random sampling to comprise the 50 vignettes of the survey. In the event of low-quality pictures that might prevent adequate evaluation ($n = 6$), the case was discarded and replaced by the next one in the random list.

2.3 | Measurement

Pediatric dermatologists (raters) were invited by email to participate in the study. On acceptance, they were sent a link to a secure web application containing a survey with 50 clinical vignettes. Each vignette contained a picture of an IH and information about age, sex, and size of the tumor. At the foot of each vignette, the question "Would you treat this patient?" was asked, and raters had to choose one of three of the following options: (1) No (observation), (2) Yes, I would treat with topical timolol, or (3) Yes, I would treat with oral propranolol.

Raters completing the survey were re-sent a new link 1 month later, for a second completion of the survey, with the vignettes randomly reordered, to allow for intra-rater agreement analyses. The clinical vignettes were the same on both occasions and for all participating dermatologists.

2.4 | Main outcome measure

Inter-rater agreement on the treatment option in the first survey was measured with Gwet's AC_1 coefficient, which is more robust than other kappa statistics.[7] For intra-rater agreement, a paired t test for testing the difference between two correlated agreement coefficients was used (weighted Gwet's AC_2).[8] The interpretation of kappa was done following the benchmark established by Landis and Koch, slightly modified by Altman[9]: a coefficient of 0.2 or less is considered "poor" agreement; 0.21–0.40: "fair" agreement; 0.41–0.60: "moderate" agreement; 0.61–0.80: "good"; and 0.81–1.00 "very good" agreement.

2.5 | Secondary outcome measure

The objective was to assess whether the inter-rater agreement varied according to the IH risk category. For this objective, the risk category was established by two pediatric dermatologists who participated in the conception of the study and were excluded from survey completion. They separately reviewed all vignettes and established an IH risk category for each case, according to the classification of the Spanish consensus on IH, based on that proposed by Luu and Frieden.[10] The Spanish consensus adds the category of "intermediate risk" for IH that is either located on skin folds, pedunculated, segmental larger than 5 cm (except on lumbar or facial area) or larger than 3 cm when located on hands.[5]

They were asked to resolve any potential disagreements on the classification (none are there). These risk categories were used for the stratified analysis of the respondents.

2.5.1 | Sample size

Lacking reference guidelines on the estimation of sample size in the case of agreement studies with more than two categories with multiple raters, simulations were performed with the R kappaSize program,[11] which allows simulations for less than seven raters. Evaluating 50 vignettes in less than 20 min proved feasible. Based on simulations, we estimated that 50 vignettes and 15 raters would give acceptable results.

2.6 | Data analysis

Study data were collected and managed using REDCap electronic data capture tools hosted at the Research Unit of the Fundación Piel Sana AEDV. Data analysis was conducted with Stata Statistical Software (version 16; StataCorp), except for sample size calculations and intra-rater analyses, for which R core Team version 4.0.4 was used (package KappaSize and the R functions for paired t test for agreement coefficients).[12]

3 | RESULTS

The final sample of clinical vignettes was composed of 50 IH with a mean age of 3.8 months (SD = 1.9), with 74% girls. They represented 7.8% of the AEDV nationwide IH cohort, and the risk categorization of the selected vignettes by two independent pediatric dermatologists was as follows: 8 high risk (16%), 15 intermediate risk (30%), and 27 low risk (54%).

Of 27 pediatric dermatologists invited to participate in this inter-rater agreement study on the treatment of IH, 24 (89%) completed the first survey, with all 23 completing both. Overall, each dermatologist showed different therapeutic preferences (Figure 1).

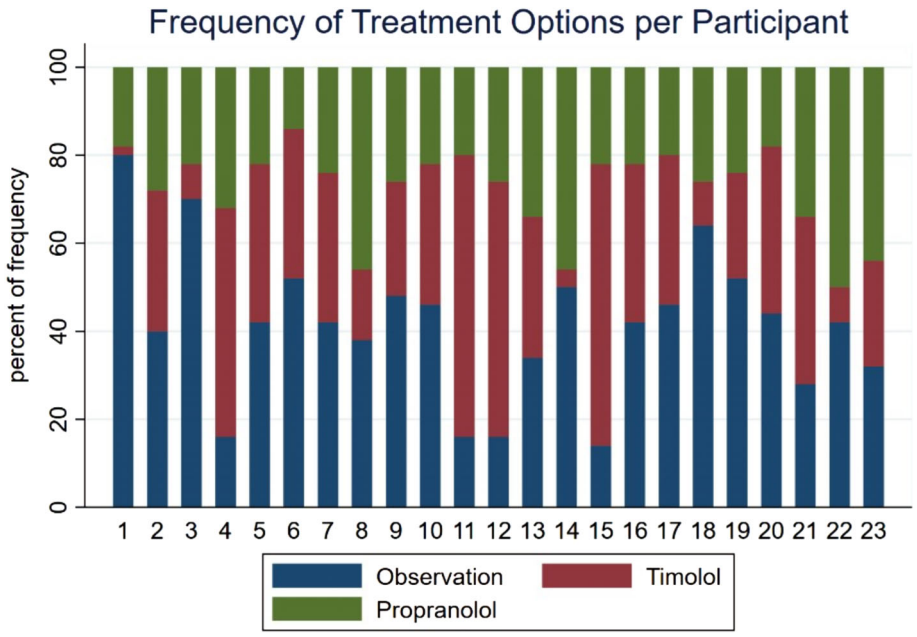

FIGURE 1 Distribution of the treatments chosen by each dermatologist. Each bar represents the percentage of therapeutic options selected by each of the 23 participating dermatologists

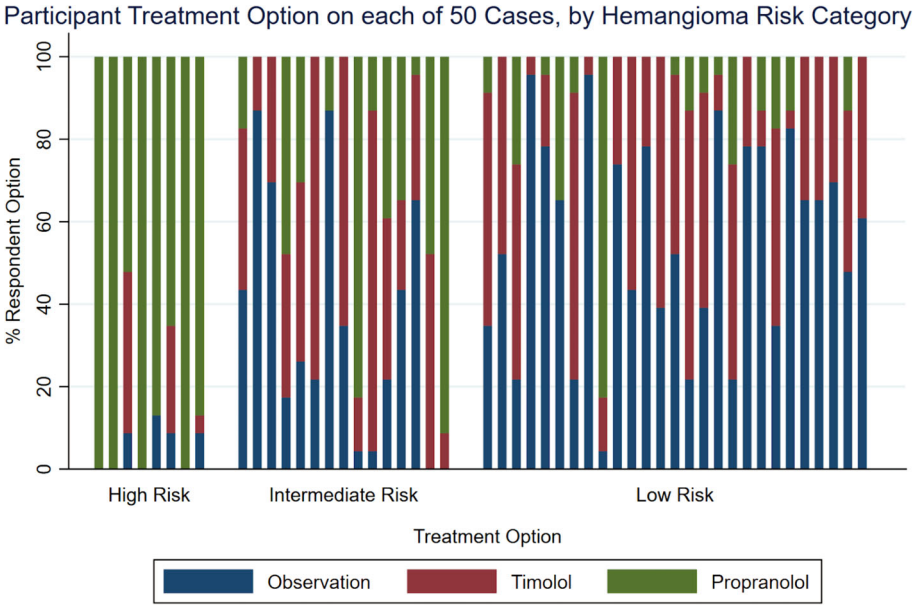

FIGURE 2 Relative frequency of the treatment option selected by 23 pediatric dermatologists in each of 50 infantile hemangioma cases, stratified by hemangioma risk category. Each bar represents a vignette and the % of each option chosen by the 23 participating dermatologists. Good agreement would be indicated by the bar being mostly of one color

TABLE 1 Frequency of the therapeutic option selected by hemangioma risk category ($N = 23$ raters, evaluating each the same 50 hemangiomas; 1150 ratings)

Hemangioma		Therapeutic option, *N* (row %)			
Risk category	***N* (col%)**	**Observation**	**Timolol**	**Propranolol**	**Total**
Low	27 (54.0)	347 (55.9)	210 (33.8)	64 (10.3)	621 (100)
Intermediate	15 (30.0)	121 (35.1)	127 (36.8)	97 (28.1)	345 (100)
High	8 (16.0)	9 (4.9)	16 (8.7)	159 (86.4)	184(100)
Total	50 (100.0)	477 (41.5)	353 (30.7)	320 (27.8)	1150 (100)

Note: Chi-squared test on four degrees of freedom = 424.5, $p < .01$.

The overall inter-rater agreement in the treatment of IH obtained for the first survey was 0.39 (95% confidence interval [CI]: 0.30–0.47), indicating fair agreement (Figure 2). When stratifying by risk category, good agreement was reached for high-risk IH ($AC_1 = 0.77$, 95% CI: 0.51–1.00), whereas for intermediate- and low-risk categories, the agreement was only fair, with similar results for both categories ($AC_1 = 0.31$, 95% CI: 0.16–0.46 and $AC_1 = 0.38$, 95% CI: 0.27–0.48, respectively). Propranolol was the main option selected for high-risk IH (86.4%, Table 1). For low-risk IH, the most frequent treatment option was observation (55.9%), whereas for intermediate-risk

IH, timolol was the first option (36.8%) closely followed by observation (35.1%). Treatment options were strongly associated with risk category type ($p < .01$, Table 1).

Regarding the intra-rater agreement, we obtained very similar results on both occasions, with strong evidence against a difference between Gwet's agreement coefficients ($p = .87$).

4 | DISCUSSION

Intra-rater agreement among Spanish pediatric dermatologists regarding the treatment of IH was very high, with dermatologists providing very similar answers on two occasions. However, the inter-observer agreement was only fair; different dermatologists offered different therapeutic options when confronted with the same vignette (File S1). This disagreement was most significant for intermediate- and low-risk IH. Variability occurred despite the existence of clinical practice guidelines that aim to minimize variation in care and could be explained by the fact that guidelines focus on the triage and treatment of high-risk IH.[4]

Previous IH consensus statements have been mainly focused on the use of propranolol.[13–16] A comparison of several guidelines shows that such recommendations are heterogeneous regarding criteria on initiation, evaluation, admission, starting dose, and monitoring.[17] In 2015, Kumar et al.[18] published a survey of 149 pediatric dermatologists, which showed that 75% of participants did not follow U.S. consensus guidelines for propranolol use[16] and that 91% prescribed topical timolol.

Our data are consistent with strong agreement among pediatric dermatologists in the treatment of high-risk hemangiomas, as the use of propranolol is the treatment of choice in this risk category (86.4% of cases). However, therapy chosen in the intermediate- and low-risk categories varied, with topical timolol elected in a third. Timolol has emerged as a treatment option in the last 10 years based on its better safety profile when compared to systemic propranolol.[2] Nevertheless, controversy remains regarding its efficacy, as it has not yet been tested in large rigorous clinical trials.[19,20] The therapeutic choice of each dermatologist in non-high-risk IH might be influenced by their own experience,[21] especially regarding the use of timolol. In the clinical setting, the decision whether to observe or use topical timolol is less straightforward and typically depends on a joint decision between the dermatologist and family after a discussion of the risk and benefit of both options.

The strengths of our study include the random selection of cases based on a cohort representative of clinical practice and the high response rate (85,2%) among a representative population of pediatric dermatologists experienced in the evaluation and treatment of IH. The data items provided in the vignettes (age, location, size, and photography of the lesion) were chosen to allow for appropriate IH classification. According to the categories of the Spanish consensus. We chose the therapeutic options of observation, topical timolol, and systemic propranolol, given their strong recommendation grade in clinical practice guidelines and consensus conferences.

The main limitations of our study are those inherent to the use of vignettes. Comparing results from our survey vignettes to those of clinical practice from the Spanish Registry,[22] we observed the election of propranolol in a higher percentage of patients in our survey (28% vs 12%). This discrepancy could lie in a greater ability to adhere to protocols in a simulated situation compared to reality.[23] Although the main variables used to categorize IH risk were presented in our survey, many other variables in real-life situations, such as contraindications to beta-blocker systemic treatment and parental preferences, with a large subjective component, impact treatment decision-making. However, the high intra-rater agreement supports that pediatric dermatologists will make consistent treatment decisions provided the same clinical scenario.

Another limitation is that the relatively small number of vignettes in each risk category prevents us from drawing definitive conclusions about the precise reasons for disagreement. For this, it would be necessary to expand the number of participating dermatologists and vignettes, especially for low- and intermediate-risk categories, where the agreement was lower. A future study with international participation could provide more generalizable conclusions.

Other possible treatment options, such as laser or corticosteroids, were not among the therapeutic options given and could have diminished the degree of agreement. However, their use is currently uncommon in Spain (2 patients treated with laser and none with corticosteroids out of 317 patients treated in the Spanish Registry) making this a lesser limitation.

Finally, a combination of treatments was not provided as an option in our survey. Some publications, with network meta-analysis methodology, have indicated that some combinations could be superior to monotherapy.[24] However, given this theoretical basis, the use of combined IH treatment in Spain is rare.[22]

5 | CONCLUSION

While each pediatric dermatologist, given the same information, was consistent in therapeutic decision-making, different dermatologists chose varied therapies for IH, especially for intermediate- and low-risk lesions. This supports the need for improving the evidence for selecting timolol versus observation in lower risk cases.

ACKNOWLEDGMENTS

We thank Eulalia Baselga, Cristina Salas Márquez, Carla Tubau, Raúl de Lucas, Cristina López Sánchez, Juan Carlos López Gutiérrez, María Dañino, Lucía Quintana, Carolina Prat, Julian Boix Vilanova, and Marina de Vega for their collaboration in the Spanish Hemangioma registry. We also thank the reviewers and the editors for their recommendations to improve the manuscript.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

María Colmenero https://orcid.org/0000-0003-3387-6271
Marta Feito-Rodríguez https://orcid.org/0000-0002-2645-8904
Marta Valdivielso-Ramos https://orcid.org/0000-0003-0879-1735
Minia Campos-Domínguez https://orcid.org/0000-0002-4638-2265
Aniza Giacaman https://orcid.org/0000-0002-0438-1184
Ignacio García-Doval https://orcid.org/0000-0002-6881-5260
Mercè Grau-Pérez https://orcid.org/0000-0002-7628-3495

REFERENCES

1. Kilcline C, Frieden IJ. Infantile hemangiomas: how common are they? A systematic review of the medical literature. *Pediatr Dermatol*. 2008; 25(2):168-173.
2. Wang X, Feng W, Zhao X, Liu Z, Dong L. The efficacy and safety of topical β-blockers in treating infantile hemangiomas: a meta-analysis including 11 randomized controlled trials. *Dermatology*. 2021;237(3): 433-443.
3. Hoeger PH, Harper JI, Baselga E, et al. Treatment of infantile haemangiomas: recommendations of a European expert group. *Eur J Pediatr*. 2015;174:855-865.
4. Krowchuk DP, Frieden IJ, Mancini AJ, et al. Clinical practice guideline for the management of infantile hemangiomas. *Pediatrics*. 2019; 143(1):e20183475.
5. Baselga Torres E, Bernabéu Wittel J, van Esso Arbolave DL, et al. Consenso español sobre el hemangioma infantil [Spanish consensus on infantile haemangioma]. *An Pediatr (Barc)*. 2016;85(5):256-265.
6. Kottner J, Audige L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *Int J Nurs Stud*. 2011;48(6):661-671.
7. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Med Res Methodol*. 2013;13:61.
8. Gwet KL. Testing the difference of correlated agreement coefficients for statistical significance. *Educ Psychol Meas*. 2016;76(4):609-637.
9. Altman DG. *Practical Statistics for Medical Research*. Chapman&Hall/CRC Press; 1991.
10. Luu M, Frieden IJ. Haemangioma: clinical course, complications and management. *Br J Dermatol*. 2013;169(1):20-30.
11. Rotondi MA, Donner A. A confidence interval approach to sample size estimation for interobserver agreement studies with multiple raters and outcomes. *J Clin Epidemiol*. 2012;65(7):778-784.
12. AgreeStat Software/Inter-Rater Reliability Analysis [Internet] Agreestat.com [cited 19 March 2022]. https://www.agreestat.com/software/default.html. Accessed April 25, 2022.
13. Janmohamed SR. Minimizing differences in treatment: expert- and evidence-based guidelines for propranolol treatment of infantile haemangiomas in the U.K. and beyond. *Br J Dermatol*. 2018;179(3): 553-554.
14. Solman L, Glover M, Beattie PE, et al. Oral propranolol in the treatment of proliferating infantile haemangiomas: British Society for Paediatric Dermatology consensus guidelines. *Br J Dermatol*. 2018; 179(3):582-589.
15. Smithson SL, Rademaker M, Adams S, et al. Consensus statement for the treatment of infantile haemangiomas with propranolol. *Australas J Dermatol*. 2017;58:155-159.
16. Drolet BA, Frommelt PC, Chamlin SL, et al. Initiation and use of propranolol for infantile hemangioma: report of a consensus conference. *Pediatrics*. 2013;131:128-140.
17. Der Sarkissian SA, Wargon O, Sebaratnam DF. International heterogeneity in admission criteria and monitoring for the initiation of propranolol in infantile hemangioma. *JAAD Int*. 2020;1(2):111-113.
18. Kumar MG, Coughlin C, Bayliss SJ. Outpatient use of oral propranolol and topical timolol for infantile hemangiomas: survey results and comparison with propranolol consensus statement guidelines. *Pediatr Dermatol*. 2015;32(2):171-179.
19. Novoa M, Baselga E, Beltran S, et al. Interventions for infantile haemangiomas of the skin. *Cochrane Database Syst Rev*. 2018;4: Cd006545.
20. Muñoz-Garza FZ, Ríos M, Roé-Crespo E, et al. Efficacy and safety of topical Timolol for the treatment of infantile hemangioma in the early proliferative stage: a randomized clinical trial. *JAMA Dermatol*. 2021; 157(5):583-587.
21. Risk: benefit analysis of drugs in practice. *Drug Ther Bull*. 1995;33: 33-35.
22. Cuenca-Barrales C, Baselga-Torres E, Del Boz-González J, et al. Baseline description of the Spanish academy of dermatology infantile haemangioma nationwide prospective cohort. Comparison of patients treated with propranolol in routine clinical practice with previous pivotal clinical trial data. *Actas Dermosifiliogr*. 2021; 112(9):806-816.
23. De Alencastro L, Locatelli I, Clair C, et al. Correlation of clinical decision-making with probability of disease: a web-based study among general practitioners. *PLoS One*. 2020;15(10):e0241210.
24. Fei Q, Lin Y, Chen X, et al. Treatments for infantile hemangioma: a systematic review and network meta-analysis. *EClinicalMedicine*. 2020;26:100506.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Colmenero M, del Boz J, Bernabeu Wittel J, et al. Inter- and intra-observer variability in the selection of therapy for infantile hemangiomas among pediatric dermatologists in Spain. *Pediatr Dermatol*. 2022;1-6. doi:10.1111/pde.15015