*Comparative Effectiveness Review Disposition of Comments Report*

**Research Review Title:** Diagnosis and Management of Infantile Hemangioma

Draft review available for public comment from June 1, 2015 to July 10, 2015.

**Research Review Citation:** Chinnadurai S, Snyder K, Sathe N, Fonnesbeck C, Morad A, Likis FE, Surawicz T, Ness G, Ficzere C, McPheeters ML. Diagnosis and Management of Infantile Hemangioma. Comparative Effectiveness Review No. 168. (Prepared by the Vanderbilt University Evidence-based Practice Center under Contract No. 290-2010-0009-I.) AHRQ Publication No.16-EHC002-EF. Rockville, MD: Agency for Healthcare Research and Quality; January 2016. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

## Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or E-mail. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

*Source: https://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=2170*
*Published Online: January 15, 2016*

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| TEP reviewer #1 | Clarity and Usability | Some of the paper's conclusions could clearly have implications in (primarily clinical) decision-making. | Thank you for your comment. We have noted potential implications in the "Clinical Implications" section. |
| Peer reviewer #1 | Clarity and Usability | This was an excellent report. The main points are clear, and the conclusions and research gaps very well presented. The likely outcome of this paper will be that topical timolol will be used because of safety profile, and that other topical dosage forms of other beta blockers, either new or repurposed, will be developed. Hopefuly new research will point to the mechanism of IH growth, and that a specific drug target will be identified. | Thank you for your comment. We hope that the report will be useful for informing clinical practice. |
| TEP reviewer #2 | Clarity and Usability | yes. Makes limitations of this clear. | Thank you for your comment. |
| TEP reviewer #5 | Clarity and Usability | Overall, the document is well written and structured, although the page numbering is confusing and should be revised to include all sections with one set of page numbers. CQs should be added to the Discussion in the larger document. | We believe that page numbering issue had to do with the line and page numbering in the PDF used for peer review. We have revised text on CQs in the discussion section of the main report. |
| TEP reviewer #5 | Clarity and Usability | As discussed earlier, the document provides support for the use of those modalities currently preferred by clinicians for the treatment of IH. However, because of the generally poor quality of the existing research on the subject, it does not provide many new insights that are likely to alter management strategies. In order to make the document more relevant to policy or practice decisions, the data supported early referral should be more heavily emphasized. | As noted, we did not explicitly review the timing of referral in this report. We hope that the section on implications for clinicians and policymakers will provide useful insights. |
| TEP reviewer #3 | Clarity/ usability | Yes well structured and organized Is useful to identify research gaps, need for comparative studies, outcomes that should be used in future studies | Thank you for your comment. |
| TEP reviewer #1 | Discussion | Yes, the implications of the major findings are clearly stated in the EC conclusion, as are the limitations. | Thank you for your comment. |
| TEP reviewer #3 | Discussion | imaging: when discussing intraspinal IH - spinal hemangioma and vertebral hemangioma are usually a misnomer due to issues with terminology, and actually venous malformations | Thank you for your comment. The study discussed here identified the lesions as IH. |
| TEP reviewer #3 | Discussion | p 26 timolol: may want to clarify that the subset of patients studied would have superficial IH, so that the audience doesnt think you can treat deep or segmental IH with topical timolol | We have noted that most studies of timolol included children with superficial lesions. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| TEP reviewer #3 | Discussion | For research gaps: imaging of IH is usually not required as it is a clinical diagnosis and imaging such as US and MRI sometimes cannot distinguish hemangiomas from vascular tumors. May want to mention biopsy as the preferred diagnostic aid when the diagnosis is in queston, rather than imaging, if the diagnosis of hemangioma versus a vascular tumor is the question | We have noted in the Discussion that when a diagnosis is in question, a tissue biopsy is the most accurate method to determine the diagnosis. |
| TEP reviewer #3 | Discussion | p 33 "steroids may be used in involution stage to clear residual IH...." - no, would not say this, based on knowledge of basic science and VEGF, steroids only helpful during proliferative phase | We have deleted this text. |
| TEP reviewer #3 | Discussion | Uniform scoring systems, clinical photographs and other measurements may be helpful with future studies | This point is addressed in the final research gap listed in the report (standardization of scoring tools); however, we have strengthened the wording of the statement. |
| Peer reviewer #3 | Discussion | As mentioned in the review, the threshold for treatment is likely decreasing in the era of beta-blocker therapy and those hemangiomas being treated now with Propranolol and timolol may not be comparable with those treated with systemic steroids 10-20 years ago. | We have emphasized in our discussion of meta-analysis findings that the estimates provide relative ranking of treatment options. Outcomes for any individual may vary depending on clinical presentation and lesion characteristics, but the meta-analyses provides an indication of which treatments are likely to be most effective overall relative to the others. |
| Peer reviewer #3 | Discussion | 18) I think it is important to mention that oral steroid therapy may be treatment of choice in children that have contra-indications to beta-blocker therapy or in the minority of children (5-20%) that have inadequate response to propranolol. | We have noted in the introduction that steroids may be used in children with contraindications to beta-blockers. |
| Peer reviewer #3 | Discussion | 19) When discussing laser therapy, important to note that PDL is only going to be effective for superficial lesions as this laser only penetrates 1.2 mm depth. PDL is not a treatment for deeper IH. | We have revised the applicability text to reflect this point. |
| Peer reviewer #1 | Discussion/ Conclusion | ES-21: Is there any reason to suspect that the female prevalence is related to cosmetic factors? | This question was not in the scope of the review. |
| Peer reviewer #1 | Discussion/ Conclusion | ES-23: The needs for future research are excellent. (RCT of beta-blockers, standardized rating systems for size/volume/ change in appearance). Would also add that basic reserarch is needed to determine the molecular mechanisms of IH formation and growth, to develop new and safer treatments for this common condition. | Thank you for your comment. We have modified the future research section to note this point. |
| Peer reviewer #1 | Discussion/ Conclusion | Also, common definitions for future research would aid in future such treatment comparisons. | We have noted in the future research section that adherence to common naming conventions will improve the research in this area. |
| Peer reviewer #1 | Discussion/ Conclusion | P32: re: Harms of corticosteroids- would also note the labeled pediatric indications for corticosteroids, and the age of the labels | This section of the report focuses on the harms of corticosteroids as reported in package insert data. Appendix H lists indications for the agents included. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| Peer reviewer #1 | Discussion/ Conclusion | P39 (bottom): patient withdrawal form prednisolone group is a key point | Thank you for your comment. |
| Peer reviewer #1 | Discussion/ Conclusion | P80: Although oral dosage forms of beta blockers are available, they are in tablets and capsules which have to be compounded for these infants. This adds some complexity to the interpretation of dose, given that pharmacy or parent may be doing the compounding. And since the liquid propranolol product was approved in 2014, and these studies all predate this approval, oral products would have been compounded, either by a pharmacy or a parent, although ref 13 does use a commercially prepared propranolol preparation. | Thank you for this comment. We have added information to the Discussion (limitations) section. |
| Peer reviewer #2 | Discussion/ Conclusion | Conclusions are correctly drawm. Research gaps wery well explicitated. | Thank you for your comment. |
| Peer reviewer #2 | Discussion/ Conclusion | The report is well structured and organised. | Thank you for your comment. |
| Peer reviewer #2 | Discussion/ Conclusion | Conclusions do not provide relevant new informations but are the logical consequence of mainly poor quality literature. Difficult to do better | Thank you for your comment. We hope the review is able to inform clinical decision-making and future research in the area. |
| TEP reviewer #5 | Discussion/ Conclusion | Again, the implications of literature demonstrating that most growth is completed by 5 months of age is underemphasized. Older literature suggests slower growth rates, and this, coupled with an expectation for involution, has led to delays in referral. The document should more clearly identify data and literature that support early referral for the IHs with the greatest risk of complications. | We have added more discussion of early referral to the report's introduction. As noted, we did not systematically review the timing of referral in this report. |
| TEP reviewer #5 | Discussion/ Conclusion | The Discussion section of the ES regarding the CQs does not state any real conclusions regarding the literature review performed to answer these questions. This is a missed opportunity for quality improvement. It is not clear why the contextual questions were omitted from the Discussion in the larger document but included in the same section in the Executive Summary. | We have added text about the CQ to the main report discussion. As noted, we did not systematically review the timing of referral in this report. |
| TEP reviewer #5 | Discussion/ Conclusion | Research gap section is particularly well written. | Thank you for your comment. |
| TEP reviewer #5 | Discussion/ Conclusion | ES-14/8-10- Doesn't the literature used for contextual questions also suggest support for a higher index of suspicion in children with periorbital IHs? | We did not find studies reporting an association between periorbital IH and occult IH. We have clarified that the statement referenced is regarding # of cutaneous IH and occult IH. |
| TEP reviewer #5 | Discussion/ Conclusion | ES-23/25-28- Neither of these is an indication for steroid use during involution. | We have revised this statement. |
| TEP reviewer #5 | Discussion/ Conclusion | ES-24/38-41- IHs for which use of Timolol was investigated were not likely problematic as this treatment is effective only in superficial lesions. | We have noted that most studies of timolol included children with superficial lesions. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| TEP reviewer #5 | Discussion/ Conclusion | 79/56-57- Not sure if there is a way to emphasize this point elsewhere in the document as well. The document does not address entrance criteria for studies discussed. This is important when drawing conclusions regarding efficacy of treatment, since, for example, intralesional steroids are considered primarily for bulky lesions and pulse dye laser and topical timolol are used primarily for superficial lesions. | Most comparative studies included children with multiple lesion types and did not report effectiveness outcomes by type of lesion.<br><br>We have also noted in the report's Conclusions that the estimates provided are most important as relative rankings of treatment effects and that clinicians and families may take factors such as lesion location, type, size, and patient values into consideration when making treatment choices ("With fairly wide confidence bounds and limited data in some areas, the relative differences among these estimates are of greater importance than the absolute effects. The estimates provide a relative ranking of anticipated rates of lesion clearance among treatment options. Families and clinicians making treatment decisions should also factor in elements such as lesion size, location, type, and number, which may affect choice of treatment modality, as well as patient/family preferences.") |
| TEP reviewer #2 | ES | I find your citing of studies without referencing within the report to be frustrating. How is the reader able to then go and look at the article/data for their own scrutiny? | The report follows AHRQ's publication standards, which limit the number of citations in the executive summary. We have noted that citations to all studies can be found in the main report. |
| TEP reviewer #2 | ES-Results | Page 19/371:<br>Imaging modalities are nearly always used in one of 4 clinical contexts: For diagnosis (is it a hemangioma or not), detection of liver IH, imaging for detection of PHACE, and imaging for spinal anomalies. Imaging modalities discussion does not include imaging studies for PHACE syndrome, the most common condition apart from hepatic hemangiomas where imaging is required. It does mention those where screening u/s was done for tethered cords. I am not sure why this was omitted but a sentence mentioning that it was not evaluated might be appropriate. | We did not identify comparative studies of imaging modalities for PHACE. |
| TEP reviewer #2 | ES-results | Similarly page 28 line 50 Propranolol topical intralesional and oral are grouped together but only oral has high level evidence – it strikes me that the others should not be in the same category with a true paucity of evidence. | The reviewer is referring to Table B summarizing the effectiveness of treatments, and specifically the row pertaining to harms of topical, oral and intralesional propranolol. We agree that harms reported were typically associated with oral propranolol, and we have revised this table. |
| TEP reviewer #2 | ES-results | Page 23<br>Error is referring to "long-pulsed PDL" – PDL is not long-pulsed – maybe you got confused with long pulsed NdYAG. PDL is typically from 0.45 to 6 milliseconds) which is quite short, not long. PDL at 585 or 595nm is used where PDL is cited and in the vast majority of recent studies, with (as mentioned epidermal cooling). | Some studies referred to PDL as either short or transitional PDL or a longer pulse, designated as "long pulse PDL." We have added text to note that "long pulse" was relative to a shorter pulse duration. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| Peer reviewer #4 | General | This review mainly assessed the effectiveness and harms of pharmacological and surgical interventions for infantile Hemangioma. Quantitative synthesis was conducted for pharmacological interventions. The major concern is that heterogeneity across studies is not adequately addressed and it is not clear whether the different studies are actually combinable. | Thank you for your review. The heterogeneity among studies is explicitly modeled in the meta-analysis, via the inclusion of a hierarchical random effect that allows the studies to be partially pooled. Partial pooling allows information to be shared among models without assuming that the suite of studies is identical. Please consult the appendix for modeling details. |
| TEP reviewer #1 | General | Yes, this report has some clinical merit. From the executive summary, though, I am not entirely clear re: the target audience. The level of detail (especially the length of the report, which seems excessive) seems most appropriate to a very specialized audience (ie academic pediatric dermatologists and other pediatric subspecialists who specialize in infantile hemangioma, IH). | We have streamlined the Executive Summary and full review and attempted to target the content more explicitly to a more general audience as well as specialists. |
| TEP reviewer #1 | General | A very important point ... the authors should be more sensitive with their use of the term "cosmetic", which is best left out of a report like this. Cosmetic interventions imply those procedures which aim to improve the appearance of normal body features and which are not essential to physical health. In contrast, in babies with IH which are in sensitive or psychosocially-significant locations, our treatments would be better considered as "restorative", in which we are attempting to restore physical function and minimize disfigurement (ie, attempting to return the patient to normal, rather than trying to achieve "supra-normal" status). This is a vital distinction to make, as we face mounting pressure and challenges from insurers to cover treatments for our patients with these (at times) very deforming tumors. | Thank you for noting this distinction. We have changed wording where appropriate throughout the Executive Summary and main report. |
| TEP reviewer #1 | General | The length of the entire paper seems a bit excessive. | We have streamlined the Executive Summary and full review. |
| Peer reviewer #1 | General | This is a truly excellent report. It is highly clinically meaningful, as many infants are diagnosed with hemangiomas and the best (most effective, least toxic) course of treatment is not clear. The key questions are clear and appropriate. Well done. I will be using the ES version to make my comments, using ES pagination. | Thank you for your comment. |
| TEP reviewer #3 | General | An extensive report evaluating the available literature on diagnosis and treatment of infantile hemangiomas | Thank you for your comment. |
| Peer reviewer #2 | General | Very hard work on a dificult topic with not easy to review literature. Overally very well conducted. Key question appropriate. | Thank you for your comment. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| TEP reviewer #4 | General | General Comments: - on ES-2 and there is still disagreement about which medication represents the best choice for initial medical management.....With your data I do not think this comment is true | We have revised this statement. |
| TEP reviewer #4 | General | Additionally there is no clear consensus as to when alternative.....ES-2 I would take this out adjunctive medications such as chemotherapeutic drugs is confusing and certainly not used presently (vincristine - not used presently) sirolimus is not a chemotherapeutic drug | We have revised this statement. |
| TEP reviewer #4 | General | ES 17 I understand the issues with US BUT Age is a big consideration for accuracy and yet this is not mentioned. Even though you state you do not discuss side effects This may be misleading. Why are we mentioning imaging if you are not discussing possible adverse events. On ES-12 it states that there are no studies addressing this | The 2 comparative studies addressing imaging that met criteria for the review did not discuss adverse effects related to imaging. We have added a sentence commenting on it in the research gaps section of the report. |
| Public reviewer #1 (American Academy of Pediatrics) | General | The relevant AAP committees have reviewed the AHRQ document and had no comments. | Thank you for your review. |
| TEP reviewer #5 | General | The report represents a relatively comprehensive review of the literature available on this topic and the quality of those studies. It provides support for the use of those modalities currently preferred by clinicians for the treatment of IH. However, because of the generally poor quality of the existing research on the subject, it does not provide new insights that are likely to alter management strategies. | Thank you for your review. |
| TEP reviewer #5 | General | The target population is explicitly defined. | Thank you for your review. |
| TEP reviewer #5 | General | The target audience is not well established. | The section on "Uses of this Evidence Report" outlines target audiences. We have added this information to the Executive Summary to make this clearer. |
| TEP reviewer #5 | General | Purpose is not well identified until Applicability section. | The introductory section on "Scope of the Review" discusses the purpose of the report. |
| TEP reviewer #5 | General | The report delineates several "key" and "contextual" questions. Some readers will be unfamiliar with this distinction, and the report should include some clarification. | We have clarified this distinction in the full report. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| TEP reviewer #5 | General | Additionally, it is not clear how questions were assigned to each of these two categories. One of the "key questions" in the initial guideline proposal was whether early specialist referral for IH resulted in avoidance of complications and better outcomes.<br><br>It is not clear why this became a "Contextual Question", but the topic seems underemphasized in the Executive Summary. Since many readers will rely primarily on the Executive Summary rather than the larger report, this may represent a missed opportunity for quality improvement in management of IH. | We developed Key and Contextual questions derived from the topic nomination with the input of a technical expert panel and key informants and informed by our understanding of the data available in the literature. Questions for which we anticipated, with the input of technical experts, data to allow for quantitative assessment of the comparative effectiveness of interventions were considered key questions. Questions not directly issues of effectiveness were considered contextual questions.<br><br>Questions were also posted to the AHRQ Effective Health Care web site for public input. The review did not attempt to address the question of early referral; however, it does address the characteristics of lesions that may prompt immediate intervention. While this report cannot make direct practice recommendations, the importance of early referral has been further stressed. |
| Peer reviewer #3 | General | This report was an enormous undertaking – I commend the authors on their efforts and taking on such an extensive review. Overall, it is a comprehensive, well-written review that highlights the many research gaps that exist in our understanding of infantile hemangiomas and their management. | Thank you for your comment. |
| Peer reviewer #3 | General | It is somewhat repetitive at times and may be able to be shortened to increase its readability. | We have attempted to streamline the text throughout. |
| Peer reviewer #1 | Introduction | Excellent, no comments. | Thank you for your comment. |
| TEP reviewer #2 | Introduction | I would suggest using the recently published ISSVA classification schema reference by Wassef et al – this talks about different types of infantile hemangiomas and also is relevant re: the confusion re: nosology (page 11 line 10)of other vascular anomalies. (See reference below) | We have added discussion of this system to the report's introduction. |
| TEP reviewer #2 | Introduction | Page 12 line 10 I disagree with the assertion that there is still disagreement re: "best choice for initial medical management" vis a vis systemic medication. The huge weight of evidence supports propranolol over corticosteroids. There may be some outliers but they are in a tiny minority. | We have revised this statement. |
| TEP reviewer #3 | Introduction | line 9 : "infection" suggest change to ulceration as infection of infantile hemangiomas is exceedingly rare | Corrected, thank you. |
| TEP reviewer #3 | Introduction | page 12 amongst dermatologists I do not think there is disagreement on which medication is the first choice to treat IH, propranolol is clearly first line | We have revised this statement. |
| TEP reviewer #5 | Introduction | 1/25- Recent literature no longer recognizes a plateau phase. | We have revised this statement. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| TEP reviewer #5 | Introduction | 1/26-27 conflicts with 14/23-24. | We have revised this statement. |
| TEP reviewer #5 | Introduction | 1/46-48- Imaging is also useful when the diagnosis is in question. | We have revised this statement. |
| TEP reviewer #5 | Introduction | 1/51-54- Also important due to risk of radiation exposure if CT chosen for imaging | We have revised the text to note the potential for radiation exposure. |
| TEP reviewer #5 | Introduction | 2/3-6- Suggest: "…however, refractory lesions that possess immediate risk for morbidity or mortality, such as hemangiomas obstructing the airway or visual axis, may require more immediate surgical intervention." | We have made this change. |
| TEP reviewer #5 | Introduction | 2/20-24- Propranolol was already available as a cardiac drug before 2014 and was already the drug of choice within 1-2 years after the serendipitous discovery of its utility for IH. | We have revised this statement. |
| TEP reviewer #5 | Introduction | 4/Figure 1- Endoscopy is not generally considered imaging. Why is echocardiography included here? | We have revised the analytic framework, and removed endoscopy and echocardiography. |
| Peer reviewer #4 | Introduction | Adequate detail to provide a good introduction of the disease and different factors and considerations for the disease. | Thank you for your comment. |
| TEP reviewer #1 | Key Questions | The key questions are clearly stated, although I disagree with placing the imaging KQ as #1, given the fact that this is clinically irrelevant in the vast majority of patients. As we discussed during the Key Informant stage of this process, imaging is really only used in a few discrete settings (significantly-atypical presentation, work-up for extracutaneous stigmata with syndromic patients ie PHACES, and hepatic evaluation when multiple lesions are present). In the remainder of patients (the vast majority), imaging plays no role. | The order of key questions is not indicative of their importance. |
| TEP reviewer #2 | Meta-analysis | Page 20<br>I do not think that intralesional propranolol should be mentioned or put in figure D; on the other hand triamcinolone is intralesional and should be mentioned in Figure D | One study included in our meta-analysis did assess intralesional propranolol. We have revised the figure to note that triamcinolone was intralesional. |
| TEP reviewer #1 | Methods | Methods seem sound, with appropriate inclusion and exclusion criteria for the most part (see question "d" below). | Thank you for your comment. |
| Peer reviewer #1 | Methods | Fig C- Final Health Outcomes: is there a causation for problems with thyroid function related to hemangioma liver involvment? | This question was not in the scope of the review. |
| TEP reviewer #3 | Methods | Agree with methods | Thank you for your comment. |
| Peer reviewer #2 | Methods | Appropriate and clearly stated search strategy. No criticisms about statistical methods. | Thank you for your comment. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| TEP reviewer #5 | Methods | Although the methodology is somewhat outside my area of expertise, it does seem sound. Inclusion and exclusion criteria are reasonable and search strategies are logical. The criteria for outcome measures and strength of evidence are clearly stated. Statistical methods seem appropriate. | Thank you for your review. |
| TEP reviewer #5 | Methods | 11/10-11- These lines should be clarified by adding the method of administration (e.g. intralesional triamcinolone and topical timolol) | We have revised this statement. |
| TEP reviewer #5 | Methods | the detail of this section seems adequate. The studies are well described, with few exceptions described below. The key messages are apparent. | Thank you for your comment. |
| Peer reviewer #3 | Methods | 11) In regards to treatment of hepatic IH – eliminating case series with less than 25 patients limits the available information on treatment of these patients. As this is quite rare, case series usually with small number of patients. | We have noted this as a limitation of the review. |
| Peer reviewer #4 | Methods | "Data Synthesis" section claimed that the quantitative synthesis used a Bayesian Latent variable model to account for the difference in thresholds when defining IH clearance, which is a nice idea. However, based on the description of the methods in Appendix D, the investigators did not exactly fit a latent variable model. | Paragraphs 3 and 4 of Appendix D discuss the latent model parameterization. |
| Peer reviewer #4 | Methods | The investigators did try to define an underlying latent response distribution (like typically seen in a probit model), with πjk defined for threshold j. But there is no evidence that this underlying distribution is used to generate an overall estimate that come from a "common" threshold. In fact, based on the current description of the mode, such distribution is not used in the analysis and it is still that counts of IH clearance from different studies using different thresholds were combined using a Bayesian hierarchical logistic regression model. That is, the issue of different cutoff points was not addressed in the analysis. The Python code and data used for Bayesian analysis should be provided to help with a clearer understanding of the actual model used in the analysis, and avoid misunderstanding. If the investigators did use a latent model, it is not reflected in the current Appendix D. | Again, the latent model is described in paragraphs 3 and 4 of Appendix D. In addition, we have made the fully-annotated model available as a public GitHub repository (https://github.com/fonnesbeck/IH_meta-analysis). Please see the IPython notebook (.ipynb suffix) in this repository to view the full model and output. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| Peer reviewer #4 | Methods | Also in this analysis, results from RCTs, prospective cohort studies and retrospective studies were simply combined without evaluating the consistency/ comparability of results among the different type of study designs, or whether the confounding was adequately addressed in the observational studies. This is further complicated by the different cutoff points and data types. It seems that the raw counts from the observational studies are used in the quantitative synthesis, without considering the patient characteristics? | Regarding the combining of outputs from studies of different designs: we conducted a sensitivity analysis comparing estimates from RCT-only designs to those from a fully-combined analysis. The only obvious differences between the two meta-analyses were that uncertainty is much higher in the RCT-only model, due to reduced number of studies, and the point estimate of Timolol moved, but there is almost no information about Timolol without cohort studies, so the point estimate is not meaningful. We reported the analysis using all studies in the review.<br><br>Regarding the cutoff values among studies: The different cutoff points and data types were addressed by the latent variable model formulation, which allows arbitrary cutoff values from each study that are combined to estimate the underlying continuous model. |
| Peer reviewer #4 | Methods | Further, for the logit transformation of VAS scale on 0-100, given the description of the method, it effectively assumes that if a VAS score is 85, there is 85% probability of IH clearance. Is this a reasonable assumption, or a leap of faith? There is no justification of this. It is not clear whether the use of VAS score is consistent across studies either. | In the opinion of clinical experts, and given the description of scales in each study, this was a reasonable assumption. Two studies in the meta-analysis used VAS scales, one on a 10-point scale and one on a 100-point scale. Methods for dealing with these multiple scales are described in Appendix D. |
| Peer reviewer #4 | Methods | The model presented in the first part of page D-1 estimates a combined probability for each arm, which is not the model used in the meta-analysis to estimate treatment effect. It does not consider the randomization for RCT or groups in cohort studies, and not really relevant to this analysis. I would exclude the explanation of the first model. | Appendix D provides a rationale for the model that we used, therefore we would prefer to retain it. |
| Peer reviewer #4 | Methods | For the Bayesian analysis, diagnostic of convergence? | The Gelman-Rubin diagnostic was used. The appendix has been edited to include this. |
| Peer reviewer #4 | Methods | For the estimands of interest, the expected proportion of clearance of each intervention helps to provide a measure that make the results more interpretable, but it is not a measure for comparative effectiveness. | This measure can be used comparatively by considering the difference in expected clearance. We also provide the probability of each treatment being the best, where "best" is defined as having the largest effect size. These are directly interpretable as probabilities, which must sum to one over the treatments in the set. I have also added SUCRA calculations and plots (Salanti et al JCE 2011) which is a recommended method for CE in a network analysis setting. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| TEP reviewer #1 | Results | Although I have not reviewed the complete paper and reference list, I feel that the exclusion of studies with under 25 subjects may have eliminated some important observations for rare presentations. For instance, there are several small reported series' of infants with symptomatic hepatic hemangiomas, often with associated cardiac failure and/or acquired hypothyroidism, who were treated with propranolol with beneficial effects (via improvement noted on imaging, and decreased reliance of their cardiac and thyroid medications). Large series are highly unlikely with such a rare presentation, so this valuable observational data is missed in the current analysis. | We included case series with at least 25 individuals for harms data only. As this review was focused on the comparative effectiveness of treatments, we focused on those studies that included an experimental group and a control group. We have noted our exclusion of case series as a limitation of the review process and have noted that some case series have examined rare presentations such as liver IH. |
| Peer reviewer #1 | Results | Figure 9: must add route of administration and dosage form for timolol (oral, ophalmic gel and solution) in text and in the figures and whether it was compounded. | We are not clear which figure is being referenced here, but we have noted the route of administration in all relevant tables and noted compounding as a factor adding to the complexity of the literature in the Limitations section. |
| Peer reviewer #1 | Results | For oral propranolol, prior to FDA approval products were likely also compounded so this should be noted. Is there any rationale for the chosen doses? | We only extracted dosage amount and route of administration from the included studies. |
| Peer reviewer #1 | Results | ES-12: in section on Oral propranolol vs other beta-clockers or dosage forms: “More children receiving oral propranolol… intralesional …” --intralesional propranolol? | We have added "propranolol" to clarify. |
| Peer reviewer #1 | Results | ES-15: It would be interresting to know if mothers had received antenatal betamethasone or if they had undergoen chorionic villous sampling, although I assume this information was not available. | Collecting these data was not in the scope of this review. |
| TEP reviewer #3 | Results | Yes amount of detail is appropriate | Thank you for your comment. |
| TEP reviewer #3 | Results | Key messages on effectiveness and harms are explicit figure and tables are adequate | Thank you for your comment. |
| TEP reviewer #3 | Results | Cannot think of other studies to include except there is a french study on vincristine that would be helpful to discussed second and third line treatments, however you excluded foreign language papers: Arch Pediatr. 2004 Feb;11(2):99-107.[Vincristine treatment for function- and life-threatening infantile hemangioma]. [Article in French] Enjolras O1, Brevière GM, Roger G, Tovi M, Pellegrino B, Varotti E, Soupre V, Picard A, Leverger G. | As noted, we excluded non-English studies. We have noted this as a limitation of the review. |
| Peer reviewer #2 | Results | Key messages clearly addressed. No major comments. | Thank you for your comment. |
| TEP reviewer #5 | Results | 14/30- What do we mean by "resolve"? Elsewhere, it is suggested that little involution occurs after age 4. | We have revised this statement. |
| TEP reviewer #5 | Results | 15/9- IHs are vascular lesions. They cannot outgrow their vascular supply. | We have revised this statement. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| TEP reviewer #5 | Results | 15/26-27- Suggest changing to "The finding of a subglottic hemangioma has been shown to increase with increasing cutaneous involvement in the beard distribution." | We have added this statement. |
| TEP reviewer #5 | Results | 21/19- "imaging" | Corrected, thank you. |
| TEP reviewer #5 | Results | 22/41-43- These line should be clarified by adding the method of administration (e.g. intralesional triamcinolone and topical timolol) | We have added the dosage form. |
| TEP reviewer #5 | Results | 33/23-24- Only a single case series described 100% rate of a single harm (hypotension, not supported by other literature) and another outlier described the 50% rate of sleep disturbance. It seems to me the SOE should be considered "High" based on the preponderance of evidence. | We feel the SOE for the association of propranolol with clinically important and minor harms is justified by the body of evidence. |
| TEP reviewer #5 | Results | 73/36-39- This sentence should emphasize that, although the SOE is moderate, the incidence (40/2541) is very low at 1.6% (perhaps just by including the percentage). | We have added the percentage. |
| Peer reviewer #3 | Results | My biggest concern with this review is the way in which the treatment modalities are analyzed. It gives very specific clearance rates determined through the network meta-analysis that I don't believe accurately reflect the efficacy of these medications:<br>a. Propranolol – 95% clearance<br>b. Timolol – 64% clearance<br>c. intralesionalTriamcinolone – 54%<br>d. Prednisone – 29%<br>These numbers are mentioned MANY times throughout the paper and I believe they are very misleading. The explanation of how these numbers were reached is also quite confusing.<br><br>The types of infantile hemangiomas being treated with these four modalities are very different and these treatments are also being used and reported by various subspecialists with different biases. As the paper clearly points out, there is considerable heterogeneity in the way these modalities are being used in regards to age of initiation, duration of therapy, dosing, and the degree to which safety monitoring is performed (or not). | We have added more detail on the meta-analysis and the rationale for combining studies. We have noted in the report's Discussion section that few studies reported outcomes by specific lesion types or for lesions in specific locations (i.e., most studies included multiple lesion types in multiple anatomic locations and reported aggregate results).<br><br>We have also noted in the report's Conclusions that the estimates provided are most important as relative rankings of treatment effects and that clinicians and families may take factors such as lesion location, type, size, and patient values into consideration when making treatment choices ("With fairly wide confidence bounds and limited data in some areas, the relative differences among these estimates are of greater importance than the absolute effects. The estimates provide a relative ranking of anticipated rates of lesion clearance among treatment options. Families and clinicians making treatment decisions should also factor in elements such as lesion size, location, type, and number, which may affect choice of treatment modality, as well as patient/family preferences.") |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| Peer reviewer #3 | Results | For example, the majority of patients treated in the Timolol studies have smaller and more superficial IH. Most of these patients likely have focal IH. These lesions present earlier and have the majority of their growth in the first 3-4 months of life. They have a much better prognosis overall and involute spontaneously at a younger age and often more completely than mixed or deep IH and especially larger segmental IH. The type of hemangioma being treated in the Timolol studies is very different the typical infantile hemangioma that was treated with oral steroids prior to discovery of beta-blockers. I am concerned that clinicians reading this report who don't have much experience/ expertise in the management of IH are going to understand this to mean that Timolol would be a safer and more effective treatment option than oral steroids – regardless of the hemangioma subtype. Timolol is unlikely to be effective in the treatment of most deep infantile hemangiomas and its use in larger lesions or ulcerated lesions is controversial. While Timolol therapy is very promising for small superficial IH, its efficacy for larger and more bulky lesions is limited at best. It is certainly not an appropriate treatment option for IH with ocular complications or airway/other visceral involvement. | We include information on type of lesion in each study in the summary tables. Most comparative studies included children with multiple lesion types and did not present outcomes by lesion type. We have added text to the background and discussion sections of the report noting that lesion type and location will dictate treatment choices. |
| Peer reviewer #3 | Results | I am also concerned about the comparison of intralesional Triamcinolone and Prednisone. I believe that the efficacy and safety of both of these agents is largely dependent upon the type of IH being treated. | We have added text to the background and discussion sections of the report noting that lesion type and location will dictate treatment choices. |
| Peer reviewer #3 | Results | While intralesionalTriamcinolone may be an effective agent for a bulky localized IH, it is not going to be an appropriate option for extensive, segmental IH especially if primarily superficial in nature. The majority of literature regarding intralesional steroids is focused on efficacy. There is tremendous variation in the way these injections are carried out – in regards to type of steroids infected, dose administered (rarely documents weight based doses but rather focuses on the volume of medication administered), number and timing of doses, use of concomitant therapies, etc. In the vast majority of these retrospective studies, there is no specific monitoring performed for potential adverse effects (ie blood pressure monitoring, growth monitoring, assessment of HPA axis, immunosuppression, etc). | We have noted that lesion size, type, and location may dictate treatment choices; however, the studies included in the review typically included multiple lesion types in multiple locations, and few provided outcome data by type or location. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| Peer reviewer #3 | Results | Another concern I have regarding the discussion of intralesionalTriamcinalone is that there is no mention of the additional morbidity associated when these injections are performed under anesthesia in the young infant which is often the case, especially when treating peri-orbital IH with intralesional steroids. | This was not a harm reported in the studies meeting our criteria. |
| Peer reviewer #3 | Results | I don't think one can accurately compare the treatment efficacy between these modalities that are being used during different timeframes on potentially very different types of hemangiomas that may carry very different prognoses. Certainly, there will be better treatment outcomes when evaluating effects of timolol therapy in small superficial IH that are often improving spontaneously by 6 months of age as compared with the disfiguring and/or function-threatening IH (ocular, airway, parotid, etc) that have been treated historically with oral steroids. | We have emphasized in our discussion of meta-analysis findings that the estimates provide relative ranking of treatment options. Outcomes for any individual may vary depending on clinical presentation and lesion characteristics, but the meta-analyses provides an indication of which treatments are likely to be most effective overall relative to the others. |
| Peer reviewer #3 | Results | Other comments:<br>1) I do not think going back to 1966 is a good idea for CQ 1 & 2. Would go back only to 1982 as done with the KQ 1-4. Literature prior to 1982 even more likely to include other vascular anomalies under the nomenclature of "hemangioma" | We felt it was important to capture older literature that may report information on untreated lesions. We attempted to ascertain that the lesions described were true IH but acknowledge that that determination is limited by the reporting in each study. We have noted this in the report's limitations section as well. |
| Peer reviewer #3 | Results | 2) In regards to KQ1, I do not think the aim of the study by Drolet and colleagues was to evaluate effectiveness/ harm of different imaging modalities but rather was focusing on determining the risk of underlying spinal/spinal cord issues in patients with cutaneous IH involving the lumbosacral region. | We agree but note that the study does provide some comparative data. |
| Peer reviewer #3 | Results | 3) Page 19 of 371 (page ES-9) – Would make sure to include "Intralesional" when discussing Triamcinolone. Would be nice to summarize what concentrations/ doses/number of treatments that were used in the intralesional steroid studies that were reviewed | We have added intralesional where appropriate throughout. |
| Peer reviewer #3 | Results | 4) Would define what is meant by "long pulse PDL" as referred to on page ES-13. | We have defined this--"long" is in relation to a shorter pulse duration in the comparison group. |
| Peer reviewer #3 | Results | 5) When discussing harms of Laser therapy – must include potential dangers/risks of general anesthesia if used during laser therapy | We reported harms as reported in each study, but have noted in the Introduction that harms are associated with anesthesia required for laser treatment. |
| Peer reviewer #3 | Results | 6) Page ES-15 – in regards to Corticosteroid efficacy, this paragraph reports clearance rate of "65%" with intralesionalTriamcinolone– whereas estimated at 53% throughout the rest of the review | Corrected, thank you. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| Peer reviewer #3 | Results | 7) Several studies included patients of adult age—these studies likely include patients with vascular malformations rather than true infantile hemangiomas (ex. Page ES-21 mentions a study in the second paragraph in which the individuals were between 1 month and 43 years of age – this likely included patients with different vascular anomalies, not specifically infantile hemangiomas) | We retained study with mixed populations (adult and child) if the mean age of participants was under age 18. The mean age in this study as 2.11 years. |
| Peer reviewer #3 | Results | 8) Page ES-21, paragraph 5 – would specify which medication in the sentence: "Doses over 2 mg/kg/day are not typically administered and may limit applicability of findings of two studies" -- ? pertaining to Propranolol dosing or other beta blockers? | We have noted that this statement pertains to propranolol studies. |
| Peer reviewer #3 | Results | 9) ES-22 – may also want to mention that the long-term side effects of propranolol therapy on the central nervous system -- cognition, memory, etc – is another area needing more study | We have added this statement. |
| Peer reviewer #3 | Results | 10) Agree with the research gaps identified by authors in this review | Thank you for your comment. |
| Peer reviewer #3 | Results | 12) Page 14 (page 51 of 371) – In regards to the natural history of IH, would change wording pertaining to Segmental IH growth – rather than having "later growth", would state that they may have more prolonged growth (or could say that the proliferative phase may be more prolonged/last longer). | We have changed this statement, thank you. |
| Peer reviewer #3 | Results | Also on this page – disagree with sentence that states that Involution noted to "cease around 3.5 years of age" – this is in contrast to the next paragraph which says that most lesions "resolve by age 5 to 7". There is great variation among children with IH in the time it takes for involution to be complete. Many children have involution beyond 3.5 years of age. Regression is also discussed on page 27 (64 of 371) and discussed most children have involution between the ages of 5 and 9 years. | We have noted that timing of changes in growth and involution of IH varies widely. |
| Peer reviewer #3 | Results | 13) Page 15 (52 of 371) – in the first paragraph regarding ulceration, would reword the sentence that states that "ulceration typically occurs later in the proliferation phase . . . " Ulceration may occur throughout the proliferative phase and often occurs early. It also occurs commonly in large mixed focal IH, especially in high risk anatomic locations. | We have revised this statement. |
| Peer reviewer #3 | Results | 14) Most clinicians consider LUMBAR, PELVIS, and SACRAL syndrome to be similar entities rather than 3 different syndromes as wording in this paragraph suggests | We have noted that these entities may be related. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| Peer reviewer #3 | Results | 15) Study by Jalil et al comparing intralesional Triamcinolone and Prednisolone – treatment dosage for the prednisolone is extremely low in the comparative trial (only 2 mg/kg – given every other day rather than 2-4 mg/kg daily which is usually standard dosing)—page 26 (63 of 371) | We have noted that this was a lower dose in the text. |
| Peer reviewer #3 | Results | 16) Page 33 (70 of 371) quotes oral steroids clearance rate of 25% -- reported as 29% elsewhere in the paper. | Corrected, thank you. |
| Peer reviewer #3 | Results | 17) Regarding the Timolol studies, no harms were reported but most of these did not do any specific monitoring for adverse effects. | We have noted this in the harms of beta-blockers section. |
| Peer reviewer #4 | Results | Page 14, lines 52-53: for OR = 1.05, it does not represent a 5 percent increased rate unless the event rate is rare (at least < 10%). | We have clarified this statement as reported in the study referenced. |
| Peer reviewer #4 | Results | CQ2 Page 16, the results could be more focused and synthesized to address CQ2 (What is the evidence that five or more cutaneous hemangiomas are associated with an increased risk of occult hemangiomas?), in addition to listing results for each study. | This is a contextual question for which we did not attempt quantitative synthesis but rather a narrative summary. We have attempted to streamline the presentation of rhe results nonetheless. |
| Peer reviewer #4 | Results | The results section mentioned “a network meta-analysis” in many places – but Appendix D does not describe a network analysis, just comparisons to the control group. | Multi-intervention meta-analysis is a synonymous term for network MA. We have clarified this in Appendix D. |
| Peer reviewer #4 | Results | Page 20, Agreement between US and MRI, is 0.27 a proportion or something else? If proportion, should have used a method to calculate the 95% CI that it does not include 0. Also testing whether agreement =0 is different from consistent with chance. If not a proportion, clarify what 0.27 is. | This is the proportion and CI reported in the study. |
| Peer reviewer #4 | Results | Page 21, lines 7 and 8 -- there are only 5 intraspinal IH altogether? Report the total number of patients with the disease, which will give a realistic sense on the evidence base for sensitivity. | We note the denominator in this paragraph. |
| Peer reviewer #4 | Results | Table 4, to definitely show superiority of beta-blockers, still need a measure to directly compare betablockers vs. steroids. | We submit that the effectiveness of the treatments are directly comparable because they are all modeled on the same scale. Therefore, the difference between any two estimates is the comparative effectiveness. However, we have revised the text on results to emphasize the relative differences between interventions. |
| Peer reviewer #4 | Results | Figure 5, as implied in the discussion above, it is not clear how clearance was defined across different studies. | Paragraphs 6 and 7 in Appendix D describe how the few studies that did not report outcomes as threshold counts reported their results. |
| Peer reviewer #4 | Results | Figure 6, note that estimates of sigma are typically smaller for smaller effect size | We have added a comment to this effect in the tex describing the meta-analysist. |
| Peer reviewer #4 | Results | Page 22, last paragraph, not clear that the final model used common variance or the different variance? | This is discussed in detail in Appendix D. |

| Commentator and Affiliation | Section | Comment | Response |
|---|---|---|---|
| Peer reviewer #4 | Results | To get a practical sense of heterogeneity among studies, it would still be helpful to show the clearance rate or effect size across studies in a plot, whenever the data are available. | We have added a figure outlining these data. |
| Peer reviewer #4 | Results | Across the results section, the presentation of P-values is all over the place like P<0.005, P≤0.003, P≤0.07, P≤01. On page 49, there is p values ≤ 0.01, p = 0.01, p < 0.01 – report the exact P-value or use a more consistent system. | We report the p value as reported in each study. |
| Peer reviewer #4 | Results | Page 33, lines 38-39, report the denominator. | We have added the denominator (6/10 in each arm). |
| Peer reviewer #4 | Results | Page 34, last paragraph, are the conflicting findings still combinable? | In the opinion of our clinical experts, these studies were combinable in the meta-analysis. |
| Peer reviewer #4 | Results | Table 21, 24, instead of reporting range of AE rates, may be helpful to quantitatively summarize the rates when there are multiple studies. | We have retained the range reporting of these data. |
| Peer reviewer #4 | Results | How does study quality affect the quantitative synthesis results? | We ran the analysis without poor quality studies. Results did not change substantially; thus, we have retained poor quality studies and report results for all studies. |