

Therapeutic Index (TIX) for intranasal corticosteroids in the treatment of allergic rhinitis*

T. Schäfer¹, M. Schnoor², M. Wagenmann³, L. Klimek⁴, C. Bachert⁵

1 Practice for Dermatology and Allergy, Immenstadt, Germany

2 Department of Internal Medicine, Division of Respiratory Medicine, Philipps-University Marburg, Marburg, Germany

3 ENT Department, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany

4 Center for Rhinology and Allergy, Wiesbaden, Germany

5 Upper Airways Research Laboratory, Department of Otorhinolaryngology, University Hospital Ghent, Ghent, Belgium

SUMMARY

Background: Intranasal corticosteroids (INS) are the first line treatment for allergic rhinitis (AR). To guide clinical decision-making, we created a therapeutic index (TIX) for INS reflecting efficacy and safety.

Methods: A Medline search (1966 to June 2009) was carried out to identify all placebo-controlled randomized trials, and observational reports for safety issues, with Dexamethasone, Budesonide (BUD), Fluticasone propionate (FP), Fluticasone furoate (FF), Flunisolide, Mometasone furoate (MF), Triamcinolone (TRIAM), and Beclomethasone dipropionate (BDP) as treatment for AR.

Data on three efficacy (nasal symptoms, ocular symptoms, global assessment) and three safety outcomes (epistaxis, growth, systemic ocular effects) were extracted. Meta analyses were performed for each INS and outcome and results were categorised into scores by quartiles. Scores of the three efficacy and safety outcomes were summed up to create summation scores for efficacy (ES) and side effects (AES), respectively with a maximum of 9 points. The TIX was then defined as the ratio of ES and AES.

Results: Data of 84 studies were extracted. Based on availability of data, a TIX was calculated for 6 substances. BUD showed the highest efficacy score followed by MF and TRIAM. The lowest scores for side effects were achieved by MF and TRIAM followed by FP. These findings resulted in TIX scores of 7 and 5 for MF and TRIAM, respectively, indicating a high efficacy and low potential of adverse events. Medium scores were reached by BUD and FP and lower scores by BDP and FF.

Conclusion: Although safety and efficacy is proven for all available INS by multiple studies, the systematic aggregation and analysis of data allows for a differentiated summary on clinically important features.

Key words: allergic rhinitis, efficacy, intranasal corticosteroids, safety, therapeutic index

Footnote:

Abbreviations: AES - Adverse event score; AR - Allergic rhinitis; BDP - Beclomethasone dipropionate; BUD - Budesonide; CI - Confidence interval; ES - Efficacy score; FF - Fluticasone furoate; FP - Fluticasone propionate; INS - Intranasal glucocorticosteroids; MCID - Minimal clinical important difference; MF - Mometasone furoate; OR - Odds ratio; PGA - Patients/Physicians global assessment; SMD - Standardised mean difference; STD - Standard deviation; RQLQ - Rhinitis Quality of Life Questionnaire; TIX - Therapeutic Index; TOSS - Total ocular symptom score; TNSS - Total nasal symptom score; TRIAM - Triamcinolone; VAS - Visual analogue scale

INTRODUCTION

Allergic rhinitis (AR) is a frequent allergic disease. In many Western countries, AR is the most frequent IgE-mediated allergic disease in adults with a lifetime prevalence of more than 20% ⁽¹⁾. Recent data from the worldwide ISAAC study indicated that there is a substantial increase in prevalence in most countries in young (6 - 7 years) as well as older (13 - 14 years) children ⁽²⁾. AR is characterized by sneezing, nasal itch, nasal obstruction, and rhinorrhea. In addition the condition has a significant impact on quality of life going along with sleeplessness, fatigue, irritability and poor concentration ⁽³⁾. The economic burden of this disease in Germany has been estimated by to be as high as € 240 Million annually, including health care costs and lost productivity ⁽¹⁾.

Intranasal glucocorticosteroids (INS) are the recommended first-line prescription treatment for patients with moderate-to-severe seasonal and perennial AR. The actual ARIA (Allergic Rhinitis and its impact on Asthma) guideline considers INS to be 'the most effective pharmacologic treatment of AR' ⁽⁴⁾. Several INS are approved as medication for AR. In Germany, as in many other countries, the following eight drugs are currently available for the treatment of AR: Dexamethasone, Budesonide (BUD), Fluticasone propionate (FP), Fluticasone furoate (FF), Flunisolide, Mometasone furoate (MF), Triamcinolone (TRIAM) and Beclomethasone dipropionate (BDP). Ciclesonide was left out, because it is not available in Germany; however, fewer data on Ciclesonide would have been available. For the newer substances the safety and efficacy was demonstrated by multiple randomized controlled trials (RCT).

For clinical decision-making, it would be helpful to have a systematic and summarizing overview on the efficacy and safety of these substances in a combined assessment. We therefore aimed to develop a therapeutic index (TIX) reflecting both efficacy and safety. Such an index had been developed e.g. for topical corticosteroids for the skin ⁽⁵⁾, but to our knowledge not yet for INS. The development of this TIX should be based on rigorous methods of evidence-based medicine, and thereby guide clinical decision-making.

MATERIALS AND METHODS

The TIX for each substance was designed as a ratio of a summation efficacy and a summation safety score. These subscores summarized the results for different clinical parameters. The results for each parameter were derived from the combined results of clinical studies, which were retrieved by a systematic literature search for each drug.

Definition of outcome parameters

Before data extraction, patient relevant outcomes for both efficacy and safety in the treatment of AR were defined. It was intended to define more than one and an equal number of parameters reflecting efficacy and safety.

The authors and clinical experts (CB, MW, LK) agreed on a

Table 1. Recommended dose of INS per day.

| INS | adults | children |
|-------------|------------|----------|
| Flunisolide | 200-300 µg | 150 µg |
| BDP | 200-400 µg | 200 µg |
| FP | 200 µg | 100 µg |
| FF | 110 µg | 55 µg |
| MF | 200 µg | 100 µg |
| Triam | 220 µg | 110 µg |
| BUD | 200-400 µg | - |

selection of the outcomes, which was based on the frequency of occurrence and availability of results of outcome-specific data in the studies as well as clinical importance and relevance. Concerning safety outcomes epistaxis was chosen as the major, most frequent and frequently reported short term AE. Other short-term AE such as crusting, headache, or URTI were less common and not frequently enough reported to be included as single safety parameter in the TIX. Furthermore systemic side effects either on cortisol levels or growth or systemically induced ocular symptoms were considered to be of such clinical relevance that the inclusion of these parameters in the TIX is justified.

Finally three parameters each for efficacy: Patient rated total nasal symptom score (TNSS), Patient rated total ocular symptom score (TOSS), Patient (or physicians) global assessment (PGA) and safety: Epistaxis, Long term side effects on growth or cortisol levels (observation period at least 6 month), Systemic ocular side effects such as Glaucoma or an increased ocular pressure were defined.

Data collection

A PubMed search (1966 to June 2009) was conducted to identify potentially relevant studies. The medical subjects heading terms Dexamethasone, Flunisolide, BDP, FP, FF, MF, TRIAM or BUD and AR were used to perform keyword searches of the database. This search was restricted to randomized controlled trials and publications in English, French or German language.

The resulting references were then screened based on the following criteria:

- sample size (N ≥ 100)
- intervention (INS, approved in Germany; use of recommended dosage for AR, see Table 1)
- indication (seasonal or perennial AR).
- control group (placebo)
- sufficient data on relevant outcomes reported

Due to the lack of RCT's investigating long-term effects on growth or cortisol level and systemic ocular side effects, an additional systematic search in PubMed was conducted for these outcomes considering also other study types (observational studies, case reports).

Table 2. Scores of long term side effects and systemic ocular side effects.

| Long term side effects on growth | Systemic ocular side effects |
|---|---|
| 0 = RCT's or prospective study available, no evidence of side effects | 0 = RCT's or prospective study available, no evidence of side effects |
| 1 = Case reports available, no evidence of side effects | 1 = Case reports available, no evidence of side effects or expert opinion |
| 2 = Case report available, evidence of side effects | 2 = Case report available, evidence of side effects |
| 3 = RCT's or prospective study available, evidence of side effects | 3 = RCT's or prospective study available, evidence of side effects |

Table 3. Screening and selection process of RCTs.

| N | Dexamethasone | Flunisolide | BDP | BUD | FF | FP | MF | Triam |
|--|---------------|-------------|------|-------|-----|-------|------|-------|
| Initial sample | 16 | 39 | 79 | 99 | 93 | | 56 | 41 |
| Remained after 1 st screening of title and abstract | 2 | 24 | 47 | 39 | 12 | 50 | 24 | 20 |
| Excluded in 2 nd Screening process | | | | | | | | |
| No TIX relevant data | n=0 | n=7 | n=12 | n=11 | n=1 | n=12 | n=3 | n=5 |
| Not the recommended dose | n=0 | n=0 | n=0 | n=3 | n=1 | n=4 | n=2 | n=5 |
| Non-allergic Rhinitis | n=0 | n=4 | n=3 | n=1 | n=0 | n=0 | n=1 | n=0 |
| Less than 100 patients included | n=2 | n=13 | n=22 | n=13 | n=3 | n=12 | n=6 | n=2 |
| Included studies | n=0 | n=0 | n=10 | n=11* | n=7 | n=22# | n=12 | n=8 |

*+3 studies including less than 100 patients for TOSS calculation

#+2 studies with VAS data for TOSS

Furthermore, a systematic search for additional head-to-head comparative studies was performed. We cross-checked the results of the TIX with the head-to-head comparisons and eventually adjusted scores of the TIX to avoid conflicting results.

By taking only RCTs with the same comparator (placebo) into consideration, we tried to ensure that results for identical outcomes, which here included TNSS, TOSS, PGA, and epistaxis, should be comparable between different INSs.

Screening and extraction

A 2-stage filter process applying the eligibility criteria was performed by screening titles and abstracts first and then full texts. Data were extracted by MS in a standardized way into predefined tables. The extracted information included 1st author, publication year, intervention, observation period, study group characteristics (number, age, and sex), and indication. TNSS and TOSS (mean + standard deviation if available) at the end of the intervention were recorded. Only TNSS or TOSS rated on a 4-point scale (0 - 3) including three or more symptoms were considered in the data extraction. When a patients' global assessment score was not available, the corresponding physicians' assessment was extracted. Regarding the patients' global assessment the percentage of

patients in the category of the best available outcome (e.g. complete relief, excellent response, significant or moderate improvement) was extracted. Furthermore, the percentage of patients with epistaxis was recorded. In case TIX relevant parameters were measured in the study but results not reported detailed enough, we tried to get in contact with the authors to receive the relevant information.

Statistical analyses and calculation of TIX

For TNSS and TOSS, the difference of group means at the end of the observation period was extracted together with the corresponding STD. In case the STD was not given, it was calculated as an estimate based on the average STD of all other studies. Standardized mean differences were then calculated by meta-analyses and presented together with the corresponding 95% confidence interval. For epistaxis and PGA, percentages were compared. Odds ratio and corresponding 95% confidence intervals were reported as measures of association and stability. Meta-analyses were then performed for each single INS and for the parameters TNSS, TOSS, PGA, and epistaxis. Heterogeneity among studies for every outcome was assessed by Chi² and I² tests. Fixed effect models were chosen in case no significant heterogeneity was observed ($p < 0.05$ i.e. $I^2 < 50\%$). Otherwise random-effect

models are presented. The software RevMan 4.2 was used for these analyses.

The meta-analyses results for the different INS and the parameter TNSS, TOSS, PGA, and epistaxis were ordered and then categorized into scores from 0 to 3 by using quartiles. Table 4 illustrates the categorization as well as the number of underlying patients and studies. For the efficacy parameters (TNSS, TOSS, PGA) a high score indicates a high efficacy.

The scores of long-term side effects and systemic ocular side effects were based on the highest level of evidence reflected by the study type of available studies and its results. Table 2 shows, how scores were defined for these outcomes. Characteristics and results of the studies which were considered for the assessment of long term (growth) and systemic ocular side effects are summarized in Appendices III and IV. High scores in safety parameters indicate a high potential of side effects.

The score points for the three efficacy- and the three safety-parameters were then summarized for each INS resulting in individual summations scores of 'sum efficacy' (ES) and 'sum side effects' (AES), which could range between a minimum of 0 and maximum of 9 points. A high ES would indicate a high efficacy and a high AES a high potential for side effects. The final TIX score was then calculated as the ratio of ES and AES with a theoretical maximum of 9 points indicating an optimal balance of a maximum efficacy and a minor potential of side effects.

RESULTS

Literature search

PubMed revealed 423 potentially relevant studies. In the two screening processes, 353 studies were excluded for reasons given in detail in Table 3. Finally, we included 11 RCT's for Bud, 12 RCT's for MF, 7 RCT's for FF, 22 RCT's for FP, 10 RCT's for BDP, and 8 RCT's for TRIAM. Eleven studies were not only placebo controlled, but also comprised a head-to-head comparison, so that we considered 59 RCT's in our data extraction. To make TOSS calculations possible, we included 5 additional studies. Regarding the head-to-head comparative studies we checked three further publications. To calculate the score for long-term side effects, 14 studies were considered comprising seven RCT's (a few of them included less than 100 patients), six prospective studies and one case report. One of these studies was also used in the meta-analyses. We considered four different studies for systemic ocular side effects. Overall data from 84 studies were extracted. All included studies are reported in the appendix E1. Due to the lack of RCTs covering 100 patients or more, Flunisolide and Dexamethasone had to be excluded from the TIX calculation.

Efficacy-Outcomes

Total nasal symptom scores. The meta-analyses of the TNSS revealed that MF showed the best effect (SMD -1.34, 95% CI

-1.73 to -0.94) followed by BUD (SMD -0.68, 95%CI -0.87 to -0.50), FP (SMD -0.72, 95% CI -1.03 to -0.41), BDP (SMD -0.68, 95% CI -0.90 to -0.46), TRIAM (SMD -0.59, 95%CI -0.83 to -0.36), and FF (SMD -0.49, 95% CI -0.61 to -0.38). These findings were in line with a head-to-head comparison of FP and BUD, showing that BUD was superior to FP regarding the TNSS. Transforming these results into scores from 0 to 3 by quartiles, the following distribution was achieved: MF = 3, BUD = 2, FP = 2, BDP = 2, TRIAM = 1, FF = 0). Results of all meta-analyses together with the derived TIX scores are displayed in Table 4.

Total Ocular Symptom Score. Concerning the TOSS, BUD was superior to the other INS (SMD -1.78, 95% CI -3.68 to 0.12) resulting in a TIX-Score of 3. In this case, we had to include studies with sample sizes smaller than $n = 100$ due to lack of other studies. This result was followed by MF and TRIAM, which had a similar effect (SMD -0.35, 95% CI -0.48 to -0.22; and -0.39, 95% CI -0.76 to -0.02). The effect on ocular symptoms was lower in patients treated with FF (-0.25, 95% CI -0.37 to -0.14) or BDP (-0.29, 95% CI -0.57 to -0.01), whereas FP showed no strong effect on TOSS (SMD -0.09, 95% CI -0.61 to 0.43). The latter data were based on VAS results due to lack of other results.

Patients' Global Assessment. Concerning the PGA, the meta-analyses showed that BUD had the greatest effect (OR 6.28, 95% CI 4.14 to 9.53) followed by FP (OR 3.51, 95% CI 2.62 to 4.69), TRIAM (OR 3.02, 95% CI 2.06 to 4-43), MF (OR 2.93, 95% CI 2.20 to 3.90), BDP (OR 2.35, 95% CI 1.41 to 3.90), and FF (OR 2.29, 95% CI 1.74 to 3.01). A head-to-head comparison of MF and FP, however, revealed a significant better result in the PGA for MF compared to FP. To avoid an obvious contradiction the scores for MF and FP were equalized. Finally, the findings resulted in the following scores for PGA: BUD = 3, FP = 2, TRIAM = 2, MF = 2, BDP = 1, and FF = 0.

Safety outcomes

Epistaxis. TRIAM showed the best results regarding epistaxis (OR 0.64, 95%CI 0.15 to 2.64). Treatment with MF, FP, BDP, and BUD led to moderately increased risk of epistaxis compared to the placebo groups (OR 1.46, 95% CI 1.01 to 2.12; OR 1.46, 95% CI 1.08 to 1.90; OR 1.51, 95% CI 0.98 to 2.32; OR 1.75, 95%CI 0.48 to 6.40), whereas the risk was elevated more than two times under FF (OR 2.37, 95%CI 1.62 to 3.46). The corresponding scores were: BUD = 2, FP = 1, TRIAM = 0, MF = 1, BDP = 2, and FF = 3.

Long term effect on growth or cortisol levels. One RCT investigating the effect of FF on 24-h urine cortisol (UC) excretion found that FF recipients had similar 24-h UC excretion results compared to those receiving placebo after a 12 months treatment period. Regarding FP, Allen and colleagues found no growth suppression in children treated with FP for one year compared with a placebo group. Furthermore, Banov et al. found similar morning plasma cortisol concentrations in the

Table 4. Results of meta-analyses for each parameter by substance.

| | BDP | BUD | FF | FP | MF | Triam |
|---|--------------|-------------|-------------|-------------|--------------|--------------|
| Efficacy | | | | | | |
| TNSS (SMD) | -0.68 | -0.68 | -0.49 | -0.72 | -1.34 | -0.59 |
| 95% CI | -0.90;-0.46 | -0.87;-0.50 | -0.61;-0.38 | -1.03;-0.41 | -1.73 ;-0.94 | -0.83;-0.36 |
| Included studies (N) | 3 | 8 | 6 | 6 | 8 | 5 |
| Included patients (N) | n=669 | n=1654 | n=2106 | n=1259 | n=1814 | n=678 |
| TIX-Score (cut off for quartiles: 0.56/0.68/0.88) | 2 | 2 | 0 | 2 | 3 | 1 |
| TOSS (SMD) | | | | | | |
| TOSS (SMD) | -0.29 | -1.78B | -0.25 | -0.09# | -0.35 | -0.39 |
| 95% CI | -0.57; -0.01 | -3.68;0.12 | -0.37;-0.14 | -0.61;0.43 | -0.48;-0.22 | -0.76;-0.02 |
| Included studies (N) | 1 | 3 | 4 | 2 | 1 | 1 |
| Included patients (N) | n=201 | n=163 | n=1141 | n=479 | n=983 | n=112 |
| TIX-Score (cut off for quartiles: 0.21/0.32/0.74) | 1 | 3 | 1 | 0 | 2 | 2 |
| PGA (OR) | | | | | | |
| PGA (OR) | 2.35 | 6.28 | 2.29 | 3.51 | 2.93 | 3.02 |
| 95% CI | 1.41;3.90 | 4.14;9.53 | 1.74;3.01 | 2.62;4.69 | 2.20;3.90 | 2.06;4.43 |
| Included studies (N) | 6 | 8 | 5 | 14 | 9 | 7 |
| Included patients (N) | n=1149 | n=1447 | n=1514 | n=2669 | n=2074 | n=1261 |
| TIX-Score (cut off for quartiles: 2.34/2.97/4.20) | 1 | 3 | 0 | 2 | (1)2* | 2 |
| Safety | | | | | | |
| Epistaxis (OR) | 1.51 | 1.75 | 2.37 | 1.46 | 1.46 | 0.64 |
| 95% CI | 0.98;2.32 | 0.48;6.40 | 1.62;3.46 | 1.08;1.90 | 1.01;2.12 | 0.15;2.64 |
| Included studies (N) | 7 | 5 | 7 | 16 | 7 | 5 |
| Included patients (N) | n=1575 | n=872 | n=2318 | n=3778 | n=1826 | n=716 |
| TIX-Score (cut off for quartiles: 1.26/1.49/1.91) | 2 | 2 | 3 | 1 | 1 | 0 |

*Head to head comparison Mom > Fluti_P, β includes studies n<100, # based on VAS data

Table 5. TIX-scores for each parameter, subscales and final ratio by substance.

| Parameter | BDP | BUD | FF | FP | MF | TRIAM |
|------------------------------|-------------|------------|-------------|-----------|-----------|--------------|
| TNSS | 2 | 2 | 0 | 2 | 3 | 1 |
| TOSS | 1 | 3 | 1 | 0 | 2 | 2 |
| PGA | 1 | 3 | 0 | 2 | 2 | 2 |
| Sum efficacy (ES) | 4 | 8 | 1 | 4 | 7 | 5 |
| Epistaxis | 2 | 2 | 3 | 1 | 1 | 0 |
| Long term side effects | 3 | 2 | 0 | 0 | 0 | 0 |
| Systemic ocular side effects | 2 | 0 | 0 | 1 | 0 | 1 |
| Sum side effects (AES) | 7 | 4 | 3 | 2 | 1 | 1 |
| TIX (ES/AES) | 0.57 | 2 | 0.33 | 2 | 7 | 5 |

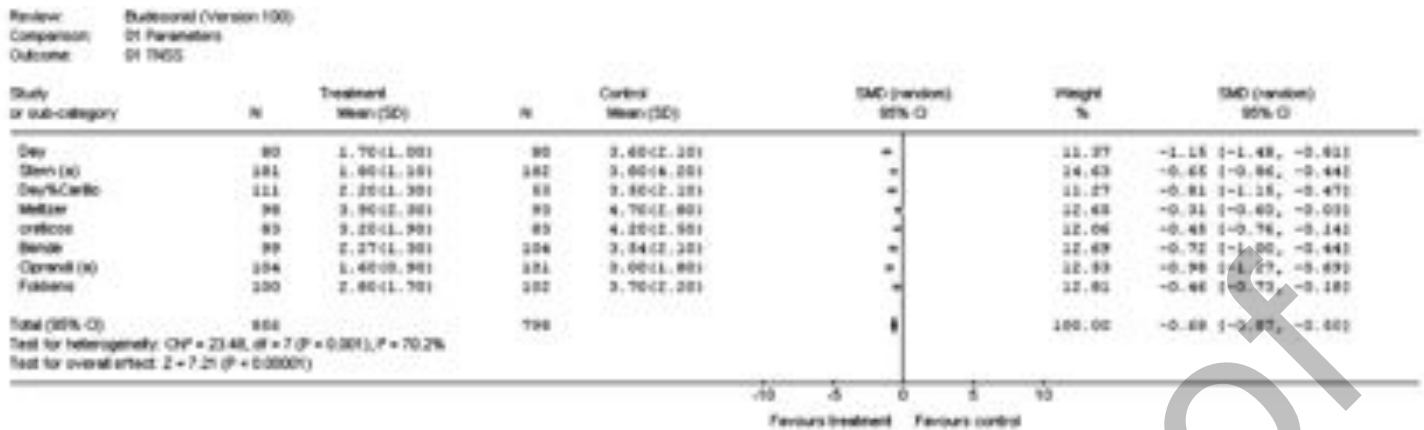


Figure 1. Example of meta analyses for the development of a therapeutic index for intranasal corticosteroids in allergic rhinitis (Budesonide, Total nasal symptom score).

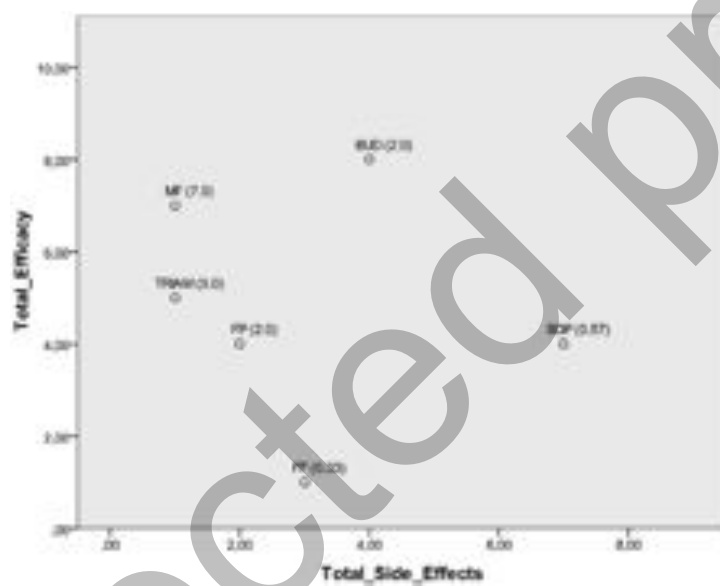


Figure 2. Therapeutic index for intranasal corticosteroids in allergic rhinitis (substance / TIX).

intervention (FP) and the placebo group after 6 months. Our literature research revealed one RCT investigating growth retardation in children with PAR after one year of treatment with MF. No significant differences in growth between the children treated with MF and the placebo group were found. In contrast, Skoner et al. found a difference in mean change in standing height after 12 month treatment with BDP, which was attributed to the treatment. Concerning TRIAM, we found two prospective studies, which did not find an effect of the treatment on statuary growth or morning serum cortisol levels. One case of growth failure after treatment with BUD was reported by Perry and colleagues. The 9 year old child received a dose of 800µg /day for 4.2 years. Following the definitions given in Table 2 these findings resulted in a TIX-score for FF=0, FP=0, MF=0, TRIAM=0, BUD=2, and BDP=3.

Ocular side effects. One RCT investigating children from 6 - 11 years showed no significant changes from baseline in intraocular pressure and posterior subcapsular cataracts after treatment

with once-daily BUD (168 µg) or MF (100 µg) for 12 months. Meltzer et al. in their RCT also did not find significant differences between children aged 6 to 11 receiving FF (110 µg) or placebo for 12 weeks regarding intraocular pressure. Opatowsky and colleagues reported two cases of increased intraocular pressure after treatment with BDP, which returned normal after discontinuation of treatment. Regarding FP or TRIAM, we did not find any study investigating the effect of INS on systemic ocular side effects. Following the definitions given in Table 2 the score for this parameter was 0 for MF, FF, and BUD, and 2 for BDP. Due to the lack of evidence and following the experts' assessment FP and TRIAM received a score of 1.

Head to head-comparisons

Overall we found 14 head-to-head comparisons, but only two studies showed significant differences between INS. The five studies comparing FP and BUD showed inconsistent results. Whereas in two studies BUD was superior to FP concern-

ing TNSS and PGA, the other three found a better effect of FP. Only one result concerning the superiority of BUD in the TNSS was statistically significant. This study also showed a non-significantly elevated frequency of epistaxis in patients treated with BUD compared to patients treated with FP. These results were in line with our TIX-score. FP and BUD have the same efficacy score of 4, but the score of epistaxis was higher in BUD compared with the score of FP.

All three studies comparing FP with BDP showed a higher efficacy of FP, but also a higher percentage of patients with epistaxis in this group. None of the differences were significant.

MF compared with FP revealed a lower TNSS after treatment with MF and a significantly higher proportion of patients with a better overall condition as assessed by PGA (55% vs. 43%, $p = 0.04$). No differences in the frequency of epistaxis were observed (17%).

The comparison of MF and BDP investigated in two studies revealed a better effect of MF regarding TNSS and PGA. One study showed a better effect in patients treated with BDP, but this effect was only marginal (TNSS 3.3 vs. 3.4). The results regarding the proportion of patients with epistaxis were inconsistent.

Only one study compared FF with BDP. In this study, the percentage of patients with substantial control of hay fever symptoms was higher in patients treated with FF than patients treated with BDP (92.3% vs. 81.0%, ns).

TIX-Score

Table 5 summarizes the scores of each parameter as well as the final TIX-score.

MF showed the highest TIX resulting from a high efficacy score and a low potential of adverse events. It was followed by TRIAM, which had an efficacy score of 5 and a low score for adverse events. FP and BUD had a similar TIX. For BUD the highest efficacy score was obtained [8], but also a moderately elevated score for adverse events [4]. FP showed a lower efficacy score, but also a lower potential of adverse events. FF and BDP, both showed a higher score for adverse events compared with the efficacy score resulting in a TIX of 0.33 and 0.57, respectively. Figure 1 illustrates the combined TIX of efficacy and side effects for each INS.

DISCUSSION

We present data on a newly developed therapeutic index for INS for the treatment of AR. Results are based on rigorous and transparent methods of evidence-based medicine including a systematic literature search and meta-analyses. With respect to single drugs Budesonide and Mometasone showed a high efficacy, whereas the potential for side effects was high for Beclomethasone. The resulting TIX showed favourable results for Mometasone and Triamcinolone.

We could not include Dexamethasone and Flunisolide in our analysis due to the lack of a sufficient number of adequately powered studies. We restricted our analysis to studies that

included at least 100 patients, as we wanted to base our results on a certain degree of statistical stability. Including smaller studies, the meta-analyses indicated a bias towards more extreme values, as expected (data not shown). We could also observe a tendency towards larger sample sizes over time, indicating that the results for substances, which results are based on older studies, are more prone to sample size error.

We did not include Ciclesonide in our analysis, because we restricted the selection of substances to those, which are approved in Germany. However, the methods used here could easily be applied to studies on any other drugs.

Summation scores for efficacy and safety were based on an equal number and more than one parameter. Having an equal number and a consistent scoring (0 - 3) assured that the ratio would equal one in case efficacy and side effects are balanced. Consequently a ratio > 1 indicates that efficacy outweighs the potential of side effects and the opposite would apply for ratios < 1 . Having included three parameters each for the safety and efficacy score a comprehensive assessment including different perspectives of efficacy and safety should be possible. Certainly there is some overlap and correlation between these parameters as e.g. nose and eye symptoms should also be reflected by the overall patients' assessment. Differences in the scores for these parameters, however, indicate that the parameters are not interchangeable and have different meanings. We selected patient-oriented and clinically relevant outcomes as we considered these to be more meaningful and appropriate than technical or surrogate parameters such as nasal air flow.

There is substantial literature on quality of life research in allergic rhinitis and several instruments for its assessment have been developed and validated ⁽⁶⁾. We would have liked to include QoL measures in the TIX, however, due to insufficient data we were unable to define this as an outcome parameter. There seems to be a trend over time for defining QoL measures as primary endpoint in recent clinical studies.

Similar, there were not enough reported data on peak (nasal) inspiratory flow measurements or rhinomanometry to have these included as an outcome parameter, although we would have considered this measure as technical surrogate parameter rather than as patient-oriented clinical endpoint.

Although differences between substances in the summary results of the parameters were found, it is difficult to assess the clinical relevance of these differences. A popular concept, which allows the interpretation of statistical results on the basis of clinical relevance is the deduction of minimal clinical important differences (MCID). Two main strategies are followed to define MCID. One is based on differences in the patient's global assessment of well-being which correspond to changes in the parameter of interest ('anchor-based'), the other is based on the STD of the measure at baseline and sets cut-off limits like 0.2 times the baseline STD ('distribution-based'). A recent study of Barnes and co-workers has applied both methods in rhinitis patients in order to establish MCID levels for a quality

of life questionnaire (Mini RQLQ), the peak nasal inspiratory flow, and the total nasal symptom score⁽⁷⁾. Whereas results for the Mini RQLQ and nasal flow are not applicable to this TIX, the MCID for the TNSS was given with 0.55 units. Applying this limit to our findings would lead to the interpretation that changes in TNSS for all INS except FF compared to placebo are clinically meaningful. Furthermore, differences between the effects of MF and all other INS except BUD and the difference between BUD and FF can be considered as clinically important.

The results of the single outcome parameters were transformed into a 4-point-scale by quartile ranges. On purpose this automatically leads to a degree of differentiation, which would not have been reached when e.g. categorization would have been made on the basis of quartiles with equal distances between the minimum and maximum values. Given that for all substances safety and efficacy is proven adequately, a differentiation for clinical purposes seems justified and there is also indication that this is clinically relevant.

There are certainly different methodological ways to develop a TIX and the results might also be different. We aimed to apply transparent, rigorous, feasible and plausible methods of evidence-based medicine to all substances in the same way. It is typical, however, for systematic secondary data analysis like this that the quality and number of the studies, overall number of patients, and assessment of the parameters differ by and within substance groups. This also goes for the duration and dosage of the treatment. We included only studies using the recommended dosages for the treatment of AR. The observation periods varied between studies, but the majority of investigations lasted between 2 and 4 weeks.

In essence, although these indirect comparisons rely on the same comparator (placebo), these methodological shortcomings should be taken into account when interpreting the results.

There is hardly any study to compare our results with. An evidence-based approach to assess efficacy and safety in a combined parameter comparing different INS was published by Portnoy and co-workers⁽⁸⁾. Numbers needed to treat (NNT) and numbers needed to harm (NNH) were calculated based on the percentage of patients who benefited compared to the placebo group i.e. the percentage of epistaxis. Taking these parameters into account a treatment threshold was calculated which was defined as the probability of a disease below which no treatment is given and above which treatment is provided considering safety and efficacy. Following that a low threshold indicates a favourable ratio of efficacy and safety. The authors studied the INS TRIAM, Fluticasone, BUD, and MF. In accordance with our results the best threshold was determined for MF (5%), whereas Fluticasone showed the highest threshold (13%). This study, however, considered only one parameter each for safety and efficacy, did not differentiate between FF and FP, did not consider BDP, and based its calculation

not on the entire available evidence, but only a single study for each substance.

The results of this TIX should help to support the daily clinical decision making process. This process, however, is also based on the personal experience of the care giver, patient preferences, and costs. We developed this TIX on clinical and not economic information. Daily treatment costs, however, are different. According to the Rote Liste® 2010, in Germany, the costs for a daily dose of BDP are lower (€ 0.32 to €0.64) than the costs for the treatment with BUD (€ 0.73) or MF (€ 0.74). The costs of FP and TRIAM are similar with € 0.79 and € 0.80, whereas the costs of FF are the highest with € 0.99 per day. Costs do also differ by country and are in general, compared to Germany, somewhat lower in the UK and higher in the USA. For a full economic evaluation appropriate cost comparative studies must be considered. This, however, was not the purpose of this project.

We conclude that although safety and efficacy is proven for all available INS by multiple studies, this systematic aggregation and analysis of data allows for a differentiated summary on clinically important features, which may help to support clinical decision making.

AUTHOR CONTRIBUTIONS

TS: conception and design of the study, screening, analysis and interpretation of the data, drafted the article, approved the submitted version of the manuscript. MS: contributed to the conception and design of the study, screening, data extraction, drafted the article, approved the submitted version of the manuscript. MW: contributed to the conception and design of the study, interpretation of the data, critically revised the article for important intellectual content and approved the submitted version of the manuscript. LK: contributed to the conception and design of the study, interpretation of the data, critically revised the article for important intellectual content and approved the submitted version of the manuscript. CB: contributed to the conception and design of the study, interpretation of the data, critically revised the article for important intellectual content and approved the submitted version of the manuscript.

CONFLICT OF INTEREST STATEMENT / FUNDING

TS: This project was funded by an unrestricted grant of Essex Pharma GmbH, Munich Germany. TS received honoraria as consultant or speaker from pharmaceutical companies incl. Essex Pharma and GSK. MS: none declared; MW: received honoraria as consultant or speaker from pharmaceutical companies incl. Essex Pharma and GSK. LK: received honoraria as consultant or speaker from pharmaceutical companies incl. Essex Pharma and GSK. CB: received honoraria as consultant or speaker from pharmaceutical companies incl. Essex Pharma, GSK, UCB, MSD, Uriach, Faes, ALK, HAL, Bencard, Stallergen.

REFERENCES

1. Ring J, Bachert C, Bauer C, Czech W. Weißbuch Allergie in Deutschland, 3. Auflage. München: Urban & Vogel, 2010.
2. Asher M, Montefort S, Björkstén B, Lai CK, Strachan D, Weiland S, Williams H, Group IPTS. Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multi-country cross-sectional surveys. *Lancet*. 2006; 368: 733-743.
3. Juniper E. Quality of life in adults and children with asthma and rhinitis. *Allergy*. 1997; 52: 971-977.
4. Bousquet J, Khaltaev N, Cruz A, et al. Allergic Rhinitis and its Impact on Asthma (ARIA) 2008 update (in collaboration with the World Health Organization, GA(2)LEN and AllerGen). *Allergy*. 2008; 63 Suppl. 86: 8-160.
5. Luger T, Loske K, Elsner P, Kapp A, Kerscher M, Korting H, et al. Topical skin therapy with glucocorticoids--therapeutic index. *J Dtsch Dermatol Ges*. 2004; 2: 629-634.
6. Kremer B. Quality of life scales in allergic rhinitis. *Curr Opin Allergy Clin Immunol*. 2004; 4: 171-176.
7. Barnes M, Vaidyanathan S, Williamson P, Lipworth B. The minimal clinically important difference in allergic rhinitis. *Clin Exp Allergy*. 2010; 40: 242-250.
8. Portnoy J, Van Osdol T, Williams P. Evidence-based strategies for treatment of allergic rhinitis. *Curr Allergy Asthma Rep*. 2004; 4: 439-446.

Prof. Dr. med. Torsten Schäfer, MPH
Practice for Dermatology and Allergy
Kemptener Str. 8
87509 Immenstadt
Germany
E-mail: info@hautarzt-immenstadt.de

Corrected proof