Supplementary Table S1. Risk of bias assessment

| Study | Random sequence generation | Allocation concealment | Blinding of participants and personnel | Blinding of outcome assessment | Incomplete outcome data | Selective reporting | Other |
|--------------------|----------------------------------|------------------------|--|--------------------------------|----------------------------|---------------------|-------|
| Bieber 2007 | Low | Unclear | Low | Low | Low | Low | Low |
| Doss 2009 | Low | Low | Low | Low | Low | Low | Low |
| Doss 2010 | Low | Low | Low | Low | Low | Low | Low |
| Hofman 2006 | Low | Low | Low | Low | Low | Low | Low |
| Luger 2001 | Low | Unclear | Low | Low | Low | Low | Low |
| Luger 2004 | Low | Unclear | Low | Low | Unclear | Low | Low |
| Mandelin 2010 | Unclear | Unclear | Low | Low | Low | Low | Low |
| Neumann 2008 | Low | Unclear | Unclear | Unclear | Unclear | Low | Low |
| Reitamo 2002a | Low | Low | Low | Low | Low | Low | Low |
| Reitamo 2002b | Low | Low | Low | Low | Low | Low | Low |
| Reitamo 2004 | Low | Unclear | Low | Low | Low | Low | Low |
| Reitamo 2005 | Low | Low | Low | Low | Low | Low | Low |
| Sigurgeirsson 2015 | Low | Unclear | Unclear | Unclear | Low | Low | Low |
| Sikder 2005 | Low | Unclear | Low | Unclear | Low | Low | Low |

Supplementary Table S2. Summary of the results. Quality of evidence assessment

| Outcome | Study | population | RR | No. of cases | Quality of evidence |
|--|--------------------------|----------------------------|------------------|----------------------|-----------------------------|
| | Calcineurin inhibitor | Topical corticosteroids | | | |
| Physician's global assessment of improvement: clear or excellent | 2713/3791 | 1976/3764 | 1.31 (1.07–1.60) | 7555 (12 studies) | Very low ^{1,2,3,4} |
| Any adverse events | 1939/2880 | 1826/2861 | 1.07 (0.86–1.33) | 5741 (9 studies) | Very low ^{1,3,4} |
| Skin burning | 757/2361 | 276/2339 | 3.24 (2.55–4.12) | 4700 (9 studies) | Moderate ^{4,5,6} |
| Pruritus | 298/2353 | 187/2324 | 1.57 (1.32–1.86) | 4677 (8 studies) | Moderate ⁴ |
| Serious adverse events | 9/1962 | 32/1932 | 0.51 (0.11–2.49) | 3894 (7 studies) | Very low ^{4,5,7} |
| Adverse events leading to discontinuation | 46/1894 | 38/1865 | 1.19 (0.77–1.83) | 3759 (7 studies) | Moderate ⁴ |

GRADE scale:

High quality: further research is very unlikely to change the confidence in the estimate of effect.

Moderate quality: further research is likely to have an important impact on the confidence in the estimate of effect and may change the estimate.

Low quality: further research is very likely to have an important impact on the confidence in the estimate of effect and is likely to change the estimate.

Very low quality: Any estimate of effect is very uncertain.

Downgraded due to the risk of bias: Allocation concealment procedure was unclear in more than one study. Downgraded due to the risk of bias: Blinding of outcome assessment was unclear in more than one study. ³Dwongraded due to inconsistency: considerable heterogeneity. ⁴Downgraded due to indirectness: corticosteroids of various potency were used. ⁵Downgraded due to inconsistency: may represent substantial heterogeneity. ⁶Upgraded due to large magnitude of effect: RR > 2. 7Downgraded due to imprecision: small number of events; small number of studies; wide CI.