Agreement and Correlation Between Different Topical Corticosteroid Potency Classification Systems

Alexa C. Bowie, MPH[1]; Mina Tadrous, PharmD, PhD[1,2,3]; Alexander Egeberg, MD, PhD, DMSc[4]; Jane Harvey PhD[5]; Stephanie J. Lax, PhD[5]; Jacob P. Thyssen MD, PhD, DMSc[4]; Aaron M. Drucker, MD, ScM[1,3,6,7]

1. Women's College Research Institute, Women's College Hospital, Toronto, Ontario, Canada; 2. Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, Ontario, Canada; 3. ICES, Toronto, Ontario, Canada; 4. Department of Dermato-Venereology & Wound Healing Centre, Bispebjerg Hospital, Copenhagen, Denmark; 5. Centre of Evidence Based Dermatology, University of Nottingham, Nottingham, UK; 6. Division of Dermatology, Department of Medicine, University of Toronto, Toronto, Ontario, Canada; 7. Division of Dermatology, Department of Medicine, Women's College Hospital, Toronto, Ontario, Canada

Corresponding Author: Aaron M. Drucker, MD, ScM, Women's College Hospital, 76 Grenville St, Toronto, ON M5S1B2, Canada (aaron.drucker@wchospital.ca).

Funding Statement: Funding for this study was obtained from the PSI Foundation and the Canadian Institutes of Health Research.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Manuscript Word Count: 1,246

Date of Revision: March 1st, 2022.

Key Points

Question: How can investigators meaningfully and transparently classify topical corticosteroid (TCS) potency for pharmacoepidemiology research?

Findings: A comprehensive list of TCS was classified using three systems and a hierarchy of sources. There was low-to-moderate agreement, but strong correlation between the potency classification systems.

Meaning: The method used to classify TCS potency in pharmacoepidemiology research could influence the results and interpretation of studies. Investigators should transparently report their TCS potency classification and consider alternative classifications in sensitivity analyses.

Abstract

Importance: Topical corticosteroids (TCS) are available in multiple potencies that influence their effectiveness and safety. Pharmacoepidemiologic studies on TCS are hampered by the absence of a universal potency classification system, limiting comparisons across studies, robust exposure classification, and clinical interpretation.

Objective: To classify TCS into three commonly used potency classification systems and evaluate the agreement and correlation between the three systems.

Design: We compiled a comprehensive list of TCS formulations using sources identified in the literature, the Ontario Drug Benefit Formulary, a recent Cochrane Review, and the Anatomical Therapeutic Classification (ATC) of the World Health Organization (WHO). TCS potency classifications were assigned and compared using the 7-category US classification system, a 4-category classification from a recent Cochrane Review largely based on the UK formulary, and the 4-category ATC classification. To facilitate comparisons across systems, we consolidated the 7-category US system into 4-categories.

Main Outcomes and Measures: We computed Cohen's weighted kappa (k_w) and Spearman rank correlation coefficients (r_s) to examine agreement and correlation between the classification systems.

Results: We included 232 unique TCS formulations (ATC: N = 231; US Classification: N = 232; Cochrane Review: N = 89). Overall, there was low-to-moderate agreement but strong correlation between the classification systems. The US classification had weak agreement with the ATC system (N = 231; κ_w 0.53, 95% confidence interval [CI], 0.45-

0.60) and moderate agreement with the Cochrane Review classification (N = 89; κ_w 0.60, 95% CI, 0.48-0.73); there was weak agreement between the ATC and Cochrane Review classifications (N = 88; κ_w 0.58 (0.46-0.71). The US classification strongly correlated with the ATC (N = 231, r_s = 0.77, 95% CI, 0.71-0.82) and Cochrane Review classification (N = 89, r_s = 0.74, 95% CI, 0.62-0.82). There was also a strong correlation between the Cochrane Review and ATC classifications (N = 88, r_s = 0.71, 95% CI, 0.58-0.80).

Conclusion and Relevance: We used multiple resources to classify 232 TCS into three potency classifications. Since these systems are often incongruent, they may yield different results in pharmacoepidemiologic studies; investigators should be transparent in their classification approach and consider alternative potency definitions in sensitivity analyses.

Background

Topical corticosteroids (TCS) are a commonly prescribed, first-line treatment for many inflammatory skin diseases. TCS have been prescribed since the 1950s demonstrating excellent effectiveness and a favorable safety record. They have been associated with predominantly local adverse effects, including skin atrophy, changes in pigmentation, steroid-induced acne, and rosacea.[1] Increasing TCS potency, application frequency, and duration of use may increase the risk of these adverse events. High-potency TCS have also been associated with extracutaneous adverse events, including adrenal suppression,[2] diabetes mellitus,[3] osteoporosis, and major osteoporotic fractures.[4] Given how commonly TCS are prescribed, and that clinician, patient, and caregiver concerns persist,[5] incorporating TCS potency into pharmacoepidemiologic studies is crucial.

TCS are available in many formulations, including different corticosteroid molecules, concentrations, and vehicles. These factors influence potency leading to differences in effectiveness and safety. Classifying TCS potency for research is hindered by the lack of a gold standard, universal classification system. If TCS potency classification systems vary and are discrepant, an investigators' choice of classification system could influence the study's results and interpretation. To our knowledge, no studies have systematically examined differences between TCS potency classification systems. To better characterize these potential discrepancies, our objective was to classify TCS formulations using three commonly used classification systems and assess their agreement and correlation.

Methods

We compiled a list of unique TCS formulations using the Ontario Drug Benefit Formulary,[6] the Anatomical Therapeutic Classification (ATC) of the World Health Organization (WHO),[7] a recent Cochrane Review,[8] and other published sources. Unique formulations were determined by the combination of corticosteroid molecule, concentration, and vehicle.

The United States Classification System is a 7-category system (1 = *Super Potent*, 2 = *Potent,* 3 = *Upper Mid-Strength,* 4 = *Mid-Strength,* 5 = *Lower Mid-Strength,* 6 = *Mild*, 7 = *Least Potent*) determined by ranking vasoconstrictive properties [9] and clinical effectiveness.[10] This system incorporates the formulation's corticosteroid molecule, concentration, and vehicle. We developed a comprehensive classification list by combining data from multiple 7-category potency lists (eFigure 1 in the Supplement). A dermatologist (AMD) and an epidemiologist (ACB) assigned previously unclassified TCS and those with incoherent classifications to appropriate potency categories (eTable 1 in the Supplement).

The ATC classifies dermatological corticosteroids (ATC Group: D07)[7] using a 4-category hierarchical classification system (1 = *Mild*, 2 = *Moderate*, 3 = *Potent*, 4 = *Very Potent*) and has been used in pharmacoepidemiologic studies.[3, 4] This system is primarily useful for drug utilization studies (rather than clinical use) and classifies TCS by their primary ingredient (e.g., betamethasone) without consideration of the salt (e.g., valerate vs. dipropionate), vehicle, or concentration.

A different 4-category classification system used in the UK (1 = *Mild*, 2 = *Moderate,* 3 = *Potent,* 4 = *Very Potent*) classifies potency by vasoconstrictive properties and concentration.[11] The earliest publication of this classification described TCS in the British National Formulary (BNF),[12] which continues to publish guidance using this system today. This, supplemented by a hierarchy of additional sources, formed the basis of a 4-category potency classification in a recent Cochrane Review on TCS for eczema.[8]

Statistical Analysis

We calculated agreement to determine the similarity of classification systems and we calculated correlation to determine their directional relationship (eMethods). Although agreement and correlation are similar concepts, they often diverge, and their relative importance differs depending on how the classification systems are being used in a given study.[13]

We reverse-coded the US classifications and consolidated them into 4-categories to analyze agreement (eTable 2 in the Supplement). We assessed agreement between the consolidated US, ATC and Cochrane Review classifications using Cohen's Kappa statistic (κ_w) with linear weighting and correlation between the reverse-coded 7-category US classification and the 4-category ATC and Cochrane Review classifications using Spearman rank correlation coefficients (r_s). All statistical analyses were conducted using SAS software (SAS OnDemand for Academics, Version 9.04).

This study followed the Guidelines for Reporting Reliability and Agreement Studies (GRRAS). Research ethics board approval was not sought for this study as it did not involve participant-level data.

Results

We classified 232 unique TCS formulations using the three classification systems (ATC: N = 231; US Classification: N = 232; Cochrane Review: N = 89) (eFigure 2 and eTable 3 in the Supplement).

The consolidated US classification had weak agreement with the ATC (N = 231; κ_w 0.53, 95% confidence interval [CI], 0.45-0.60; Table 1); 54.5% of the classifications were concordant and 37.6% US potency classifications were lower than those assigned by the ATC (Figure 1A). There was moderate agreement between the consolidated US and Cochrane Review classifications (N = 89; κ_w 0.60, 95% CI, 0.48-0.73); 64.0% of the were concordant while 33.7% of the consolidated US potency classifications were lower in potency (Figure 1B). There was weak agreement between the Cochrane Review and ATC classifications (N = 88; κ_w 0.58 (0.46-0.71); 64.0% of the classifications were concordant and 19.1% of the ATC categories were higher than the Cochrane Review (Figure 1C).

The US classification was strongly correlated with the ATC (N = 231, r_s = 0.77, 95% CI, 0.71-0.82) and Cochrane Review classifications (N = 89, r_s = 0.74, 95% CI, 0.62-0.82). The Cochrane Review and ATC classifications were also strongly correlated (N = 88, r_s = 0.71, 95% CI, 0.58-0.80).

Discussion

We used multiple sources to classify 232 TCS using three potency classification systems. The classifications showed weak-to-moderate agreement despite strong correlation, demonstrating that TCS potency classifications are inconsistent between systems. Consequently, the classification system chosen for a given pharmacoepidemiology study could substantially influence its results and clinical interpretation.

For studies on commonly prescribed medications available in different formulations, or those that are taken as needed (PRN), substituting a Defined Daily Dose (DDD) method with a more robust approach has been suggested;[14] in the context of TCS, this includes potency classifications. Given how commonly TCS are prescribed, transparent classification is necessary to strengthen dermatological outcomes and treatment. The three potency classification systems have potential strengths and weaknesses. The 7-category US system is more nuanced and may better reflect the continuous nature of TCS potency than 4-category systems, but may be difficult to interpret as the importance of a 1-level difference in classification is unclear; for example, are differences between ‘upper mid-strength’ and ‘mid-strength’ TCS clinically meaningful? The ATC classification is straightforward but overly simplistic by combining formulations with potentially different potencies. For example, in the US classification, betamethasone dipropionate 0.05% is classified as ‘Potent’ as an ointment and ‘Upper Mid-Strength’ as a cream; betamethasone valerate 0.05% is classified as ‘Mid-Strength’ as an ointment and ‘Lower Mid-Strength’ as a cream; the ATC system classifies these formulations together as ‘Potent’. Variation across classifications may limit the

comparability of findings across studies; for example, fluocinolone acetonide 0.01% cream is ‘Mild’ in the US system, ‘Moderate’ in the Cochrane Review, and ‘Potent’ in the ATC system.

This study has limitations. We determined potency from lists published in the literature, most of which did not describe how classifications were determined. Some existing classification systems, including the 5-category Japanese classification system, were not evaluated. Most classification systems were based on the vasoconstriction assay, which does not always correlate with clinical effects.[15] Future research on the clinical validity and applicability of TCS potency classification systems is needed.

Variability between classification systems suggests that research utilizing different potency classifications could yield different results. We offer a comprehensive list of TCS potency classifications to support future pharmacoepidemiology studies to strengthen evidence on TCS effectiveness and safety. Investigators should transparently report how they categorize TCS potency and consider sensitivity analyses with alternative classifications to examine the robustness of their findings.

Acknowledgement

This study received funding from the PSI Foundation and the Canadian Institutes of Health Research. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. Alexa C Bowie (Women's College Research Institute) had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Dr. Drucker reported receiving compensation from the British Journal of Dermatology (reviewer and Section Editor), the American Academy of Dermatology (guidelines writer), and the National Eczema Association (grant reviewer). Dr. Thyssen reported being an advisor for AbbVie, Almirall, Arena Pharmaceuticals, Coloplast, OM Pharma, Aslan Pharmaceuticals, Union Therapeutics, Eli Lilly & Co, LEO Pharma, Pfizer, Regeneron, and Sanofi-Genzyme, a speaker for AbbVie, Almirall, Eli Lilly & Co, LEO Pharma, Pfizer, Regeneron, and Sanofi-Genzyme, and received research grants from Pfizer, Regeneron, and Sanofi-Genzyme. Dr. Egeberg reported receiving research funding from Pfizer, Eli Lilly, Novartis, Bristol-Myers Squibb, AbbVie, Janssen Pharmaceuticals, the Danish National Psoriasis Foundation, the Simon Spies Foundation, and the Kgl Hofbundtmager Aage Bang Foundation, and honoraria as a consultant and/or speaker from AbbVie, Almirall, Leo Pharma, Zuellig Pharma Ltd., Galápagos NV, Sun Pharmaceuticals, Samsung Bioepis Co., Ltd., Pfizer, Eli Lilly and Company, Novartis, Union Therapeutics, Galderma, Dermavant, UCB, Mylan, Bristol-Myers Squibb, and Janssen Pharmaceuticals. No other disclosures were reported.

with the Same Primary Ingredient Between (A) the Consolidated US and Cochrane Review Classification Systems, (B) the Consolidated US and ATC and (C) the ATC and Cochrane Review Classification Systems.

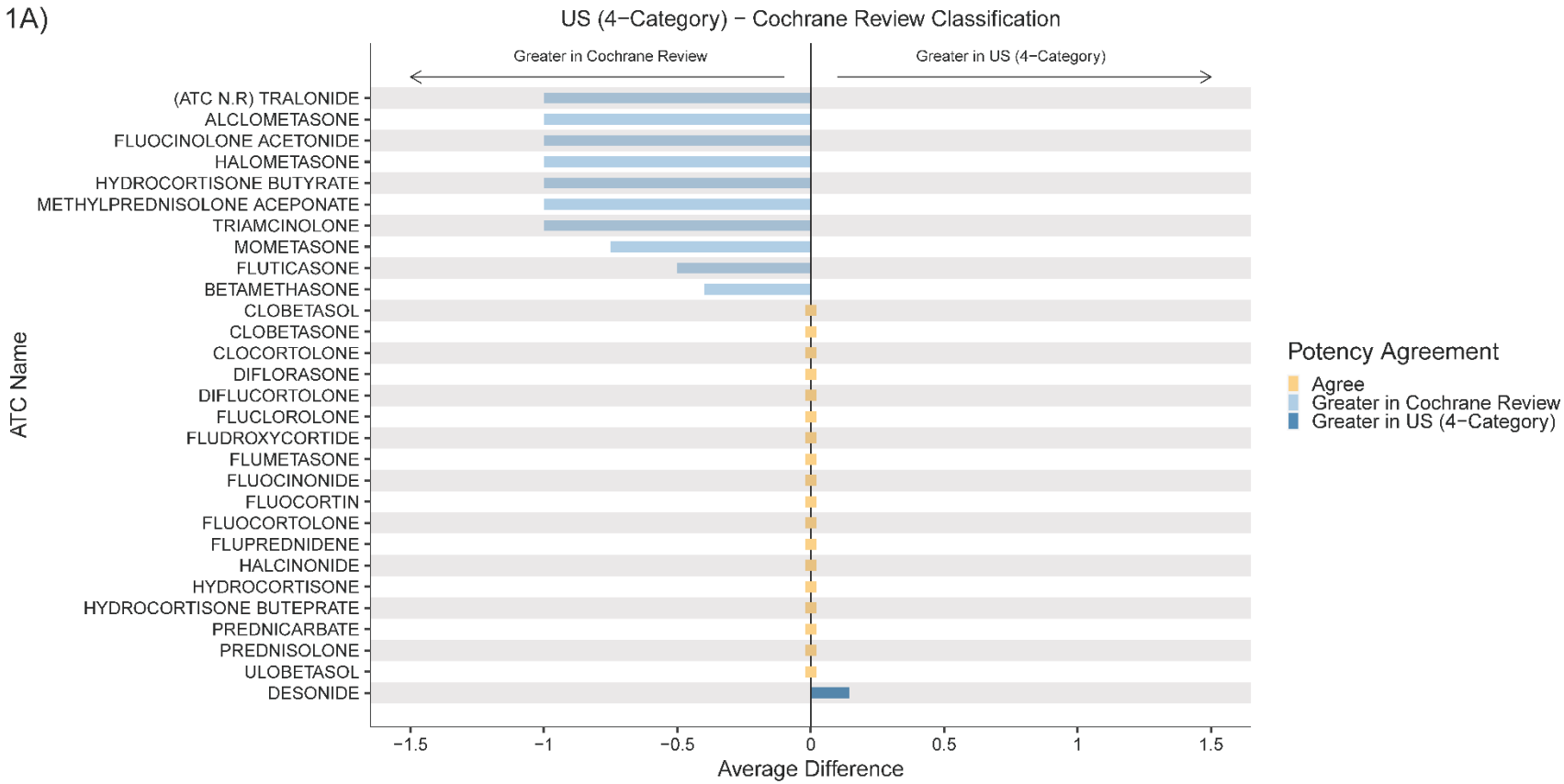

Figure 1. The Average Difference in Potency (on a 4-Category Scale) for Topical Corticosteroids with the Same Primary Ingredient Between (A) the Consolidated US and Cochrane Review Classification Systems, (B) the Consolidated US and ATC and (C) the ATC and Cochrane Review Classification Systems (continued).

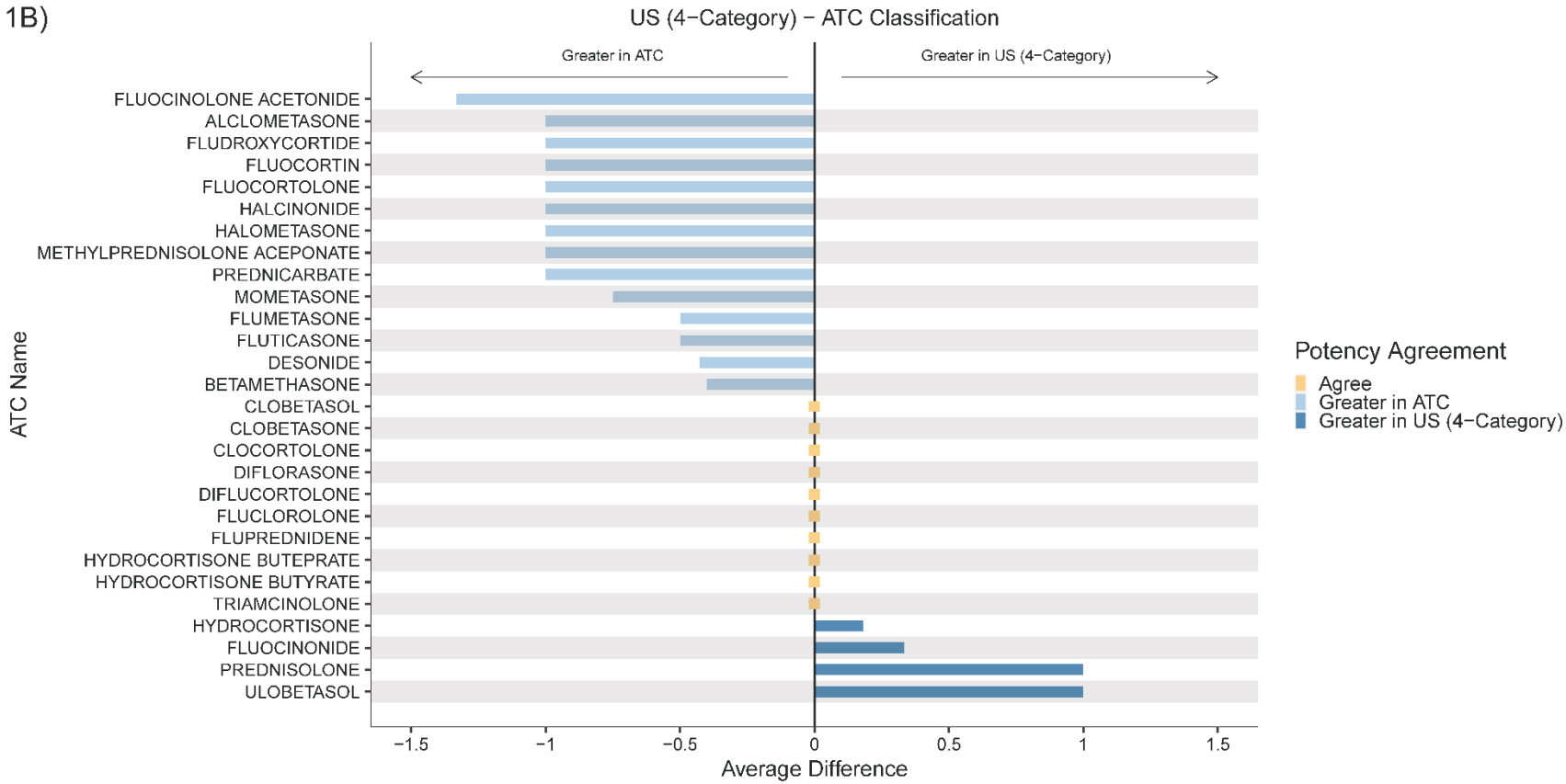

Figure 1. The Average Difference in Potency (on a 4-Category Scale) for Topical Corticosteroids with the Same Primary Ingredient Between (A) the Consolidated US and Cochrane Review Classification Systems, (B) the Consolidated US and ATC and (C) the ATC and Cochrane Review Classification Systems (continued).

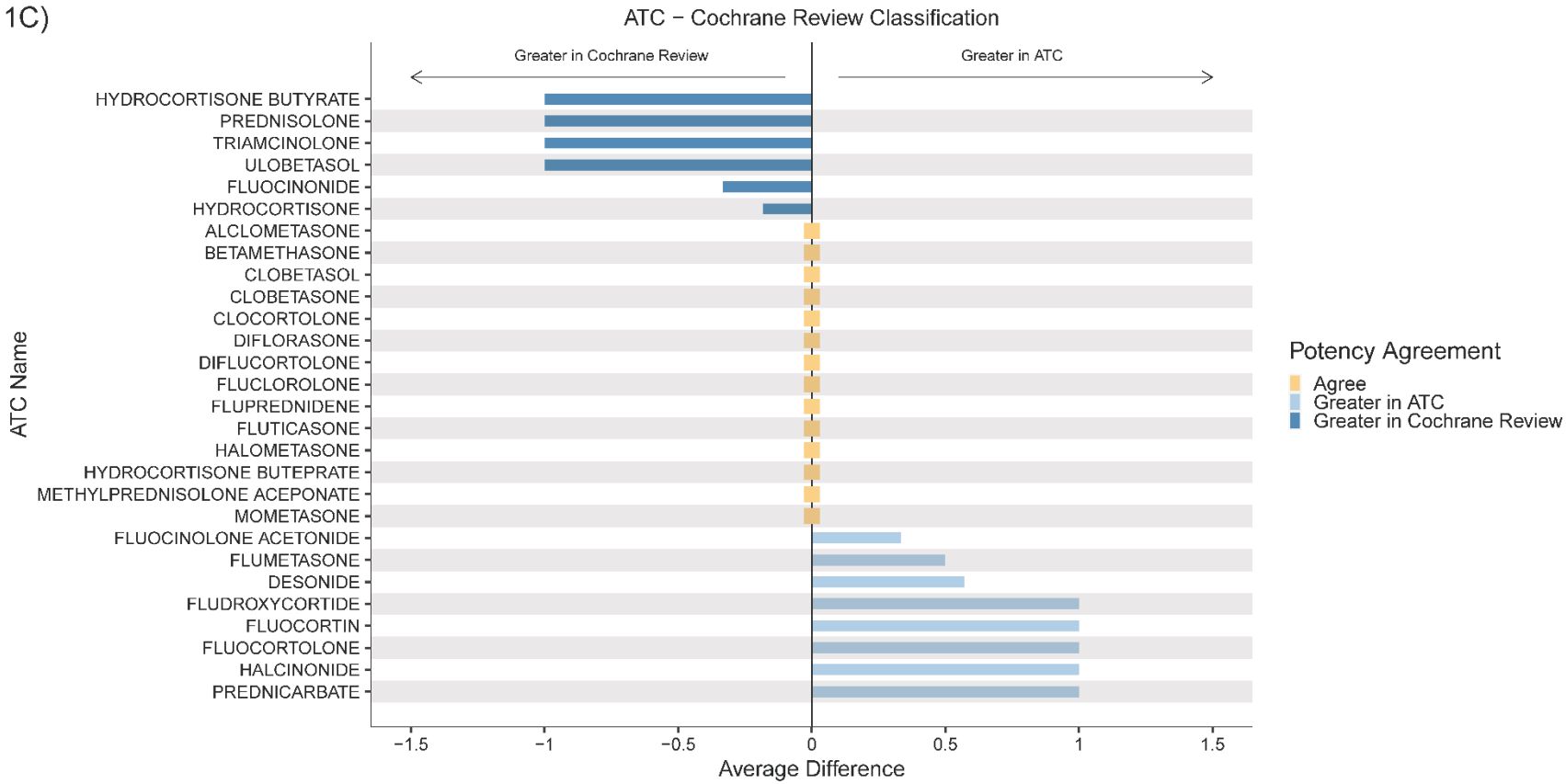

Table 1. Agreement and Correlation Between the US, ATC, and Cochrane review classification systems.

Classification System	Kappa Agreement[a] (95% CI)		Spearman Rank Correlation Coefficient (95% CI)	
	US Classification[b]	ATC	US Classification[c]	ATC
ATC	0.53 (0.45-0.60)	--	0.77 (0.71-0.82)	--
Cochrane Review	0.60 (0.48-0.73)	0.58 (0.46-0.71)	0.74 (0.62-0.82)	0.71 (0.58-0.80)

[a]Weighted using linear weighting.
[b]Consolidated (4-category).
[c]Reverse-coded (7-category).

Reference List

1. Mehta AB, Nadkarni NJ, Patil SP, Godse KV, Gautam M, Agarwal S. Topical corticosteroids in dermatology. *Indian J Dermatol Venereol Leprol*. 2016;82(4):371-378. doi:10.4103/0378-6323.178903

2. Tempark T, Phatarakijnirund V, Chatproedprai S, Watcharasindhu S, Supornsilchai V, Wananukul S. Exogenous Cushing's syndrome due to topical corticosteroid application: case report and review literature. *Endocrine*. 2010;38(3):328-334. doi:10.1007/s12020-010-9393-6

3. van der Linden MW, Penning-van Beest FJ, Nijsten T, Herings RM. Topical corticosteroids and the risk of diabetes mellitus: a nested case-control study in the Netherlands. *Drug Saf*. 2009;32(6):527-537. doi:10.2165/00002018-200932060-00008

4. Egeberg A, Schwarz P, Harsløf T, et al. Association of Potent and Very Potent Topical Corticosteroids and the Risk of Osteoporosis and Major Osteoporotic Fractures. *JAMA Dermatol*. 2021;157(3):275-282. doi:10.1001/jamadermatol.2020.4968

5. Li AW, Yin ES, Antaya RJ. Topical Corticosteroid Phobia in Atopic Dermatitis: A Systematic Review. *JAMA Dermatol*. 2017;153(10):1036-1042. doi:10.1001/jamadermatol.2017.2437

6. Ontario Drug Benefit Formulary/Comparative Drug Index. Accessed September 01, 2021. https://www.formulary.health.gov.on.ca/formulary/

7. World Health Organization. ATC/DDD Index. Accessed July 20, 2021. https://www.whocc.no/atc_ddd_index/?code=D07A&showdescription=no

8. Lax SJ, Harvey J, Axon E, et al. Strategies for using topical corticosteroids in people with eczema. *Cochrane Database Syst Rev*. In press.

9. McKenzie AW, Stoughton RB. Method for Comparing Percutaneous Absorption of Steroids. *Arch Dermatol*. 1962;86(5):608-610. doi:10.1001/archderm.1962.01590110044005

10. Cornell RC, Stoughton RB. Correlation of the vasoconstriction assay and clinical activity in psoriasis. *Arch Dermatol*. 1985;121(1):63-67. doi:10.1001/archderm.1985.01660010067020

11. Oakley R, Arents BWM, Lawton S, Danby S, Surber C. Topical corticosteroid vehicle composition and implications for clinical practice. *Clin Exp Dermatol*. 2021;46(2):259-269. doi:10.1111/ced.14473

12. Anonymous. Corticosteroids for the Skin. *Drug Ther Bull*. 1977;15(2):5-8.

13. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: Measures of agreement. *Perspect Clin Res*. 2017;8(4):187-191. doi:10.4103/picr.PICR_123_17

14. Hayes KN, Gomes T, Tadrous M. Greater than the Sum: Applying Daily-Dose Equivalents to Antipsychotic Prescription Claims to Study Real-World Effects. Brief Research Report. *Frontiers in Pharmacology*. 2021;12:709349. doi:10.3389/fphar.2021.709349

15. Humbert P, Guichard A. The topical corticosteroid classification called into question: towards a new approach. *Experimental Dermatology*. 2015;24(5)doi:https://doi.org/10.1111/exd.12677