ORIGINAL REPORT

Self-reported Versus Physician-reported Severity of Chronic Hand Eczema: Concordance Analysis Based on Data from the German Chronic Hand Eczema Patient Long-Term Management Registry

Philipp BENTZ[1], Christian APFELBACHER[2,3], Wilhelm AKST[4], Sonja MOLIN[5], Andrea BAUER[6], Peter ELSNER[7], Vera MAHLER[8], Ralph VON KIEDROWSKI[9], Jochen SCHMITT[10] and Elke WEISSHAAR[1]

[1]Department of Occupational Dermatology, Dermatology Clinic, Heidelberg University Hospital, Heidelberg, [2]Institute of Social Medicine and Health Systems Research, Otto von Guericke University Magdeburg, Magdeburg, [3]Institute for Epidemiology and Preventive Medicine/Medical Sociology, University of Regensburg, Regensburg, [4]Private practice, Prüm (Eifel), Germany, [5]Division of Dermatology, Department of Medicine, Queen's University, Kingston, Canada, [6]Department of Dermatology, University Hospital Carl Gustav Carus, Technical University Dresden, Dresden, [7]Privat practice for dermatology and allergology, SRH Hospital, Gera, [8]Paul-Ehrlich-Institut, Federal Institute for Vaccines and Biomedicines, Langen, [9]Focus practice for chronic inflammatory dermatoses, skin cancer and allergology and study center CMS GmbH, Selters (Westerwald) and [10]Centre for Evidence-based Healthcare, Medical Faculty Carl Gustav Carus, Technical University Dresden, Dresden, Germany

Self-assessment of general health status has a significant influence on patient-related outcomes. The aims of this study were to investigate and compare the level of agreement between patients' and dermatologists' assessments of the severity of chronic hand eczema. From the German registry "German Chronic Hand Eczema Patient Long-Term Management Registry" (CARPE), 1,281 pairs of patients with chronic hand eczema and their dermatologists were included. Of these, 788 pairs served as a comparison 2 years after baseline. Concordance analyses found that complete concordance between patients' and dermatologists' assessments were 16.62% at baseline and 11.47% at follow-up. Overall, patients assessed their chronic eczema at baseline as more severe than did the dermatologists; whereas, at follow-up, patients assessed their condition as less severe than the dermatologists' assessment. Bangdiwala's B showed lower values of concordance for womens' and older patients' self-assessment with the dermatologists' assessments. In conclusion, dermatologists should consider the patient's perspective and the individual's assessment of their chronic hand eczema in order to provide effective care in clinical practice.

Key words: concordance; self-reported health; hand eczema.

SIGNIFICANCE

Daily clinical practice shows that, for chronic hand eczema, self-reported and physician-reported severity may differ and may influence treatment success. This study found low levels of agreement between dermatologists' and patients' assessments of the severity of the condition. Statistical analysis showed that women and older people tended to assess the severity of their condition differently compared with the dermatologist's perspective. At follow-up patients assessed their chronic hand eczema as less severe compared with their dermatologist. To ensure better treatment success, dermatologists should take into account the personal views of patients regarding their chronic hand eczema.

Chronic hand eczema (CHE) is a complex, multifactorial dermatological disease that is challenging to diagnose and to treat (1). CHE has a severe impact on patients' quality of life and places a financial burden on patients and on society (2, 3). The exact prevalence of CHE in Germany is unknown; different studies estimate it to be in the range 2.6–16% of the population, depending on the population investigated (e.g. general population, patients' occupations) (4). Therapy for CHE comprises topical treatments, systemic treatments, ultraviolet (UV) phototherapy and the use of emollients/moisturizers (1, 5). Prevention measures and behavioural changes are also required, especially in occupational settings. Worsening or relapse of CHE can be significantly reduced by avoiding risk factors in daily life (6). Patients are thus "the primary source of their own care" (7).

Self-reported health (SRH) plays an increasingly important role in clinical studies (8). It is robust to contextual factors and short-term influences (8–10) and is derived from a complex spectrum of physical and psychological domains of health (8, 11). Thus, the patient's subjective assessment offers a valuable complement to the physician's assessment. In particular, the predictive powers of such assessment, with respect to all-cause mortality (8, 10–13) and long-term sick leave (14), have been documented.

The concordance of external, health-related factors, such as the number of physician visits and a 5-fold graded self-assessment of health status were investigated in 1980, demonstrating correlations between objective criteria and SRH (15). However, only a few other studies have addressed the concordance of SRH and physician-reported health (PRH) (9, 16–21).

The aim of this study was to analyse the concordance of assessments of severity of CHE between patients and dermatologists. This knowledge is important, since discordance could lead to impaired treatment success.

MATERIALS AND METHODS

Data source

The data for this secondary analysis was taken from the German Chronic Hand Eczema Patient Long-Term Management Registry (CARPE) (original German title: *Chronisches Handekzem-Register zum Patienten-Langzeitmanagement*). It enrolled a total of 1,281 patients between 2009 and 2016 and is the largest registry of its kind worldwide. CARPE is described in detail elsewhere (22–26). Previous analyses showed high correlations of the patients' assessments of severity of their CHE and their quality of life (22, 24, 25).

Ethics approval for the original study was obtained from the Ethics Committee of the Medical Faculty of Heidelberg (S-433/2008) and all data protection guidelines were followed in the collection, analysis and archiving of the data.

The severity of hand eczema was determined by a dermatologist, using the photoguide-based Physician Global Assessment (PGA) (23). The patient was asked to mark their self-assessment on a visual analogue scale (VAS) with the endpoints "0: Clear" and "10: Severe skin lesions", based on the question "How severe do you currently classify your hand eczema?". Transfer of the VAS results into the database was precise to the nearest decile.

To harmonize the 2 scales, patient assessments were also transformed into 5 categories. Scale values <2 correspond to the PGA degree "clear", 2–3.9 to the PGA degree "almost clear", 4–5.9 to the PGA degree "moderate", 6–7.9 to the PGA degree "severe", and 8–10 to the PGA degree "very severe". This is consistent with the approach taken in a similar study (9).

Statistical analysis

Cohen's weighted kappa ($\hat{k}_{\omega}$) is commonly used to quantify inter-rater agreement of polytomous ordinal measurement objects. However, this measure is not very reliable when inhomogeneous marginals are present (27). For this reason, Bangdiwala's weighted B ($\hat{B}_{\omega}$) was used (28). Since this method that is not a common statistical procedure, it is explained in more detail here. Simulations showed that this method is much less sensitive to inhomogeneous margins, more robust in the face of empty table rows, and overall more conservative compaired to kappa (27). To accommodate the ordinal nature of the evaluation criteria, linear Cicchetti-Allison weighting (29) was chosen to penalize deviations uniformly across the rating spectrum (30). However, as with all statistics for measuring interrater-agreement, the direction of the deviation is not taken into consideration.

In addition, Bangdiwala's B adds a visual aspect to the otherwise commonly encountered single parameter measurement of concordance, which can provide new insights on its own and is not available in conventional concordance measures (28, 30). This central element of the concordance measure developed by Bangdiwala is the "agreement chart", which is a visualization of a k*k contingency table. It consists of an area of n*n, where n is the sum of all rating processes. Within this area, k rectangles are arranged, starting from the coordinate (0,0) to (n,n), where k represents the number of individual, ordinal rating criteria. Their dimensions result from the absolute frequency of the rating criteria of both evaluating instances. Within these rectangles, concordance or deviation from the diagonal of the contingency table are represented by further rectangles with different shading. Black indicates concordance, lighter shading represents stronger deviations from the diagonal of the contingency table (31). Comparable to the kappa statistic, a single ratio for agreement can be derived, which accounts for chance agreement (28).

The "path of the rectangles" (31) also provides information on the extent to which the raters tend to prefer certain rating criteria. This can be read from the deviation of the rectangles from the diagonal of the agreement chart or their dimensions. In the case of an "optimal" inter- and intra-rating distribution, all k rectangles would be squares for which the diagonal between (0,0) and (n,n) is a mirror axis.

For a further specification of the analysis the cases were categorized by sex and age cohort.

The age groups were based on the following thresholds: 18–30 years=early professional phase, 31–50 years=central professional phase, 51–65 years=late professional phase, 66–75 years=early senior phase, 75+ years=late senior phase. Due to methodical limitations, age group 75+ years was excluded from the analysis.

Statistical analyses were performed using R 4.0.3 with the packages of the tidyverse (32) and vcd for the B statistics (33).

RESULTS

Descriptive statistics

A total of 1,281 records of severity assessment of patient-dermatologist pairs were available at baseline. At follow-up 24 months later 788 participants and their treating dermatologist remained in the cohort (follow-up survey rate 61.51%). Of these participants, 53.7% were female. Mean age was 46.98 years (standard deviation (SD) 13.55, range 18.01–84.33 years). Hand eczema had persisted for a mean of 12.78 years (SD 8.44, range 2–53 years).

After mathematically aligning the patient assessment with the PGA the ratings were cross-tabulated (**Table I**). Complete agreement between SRH and PRH in the individual categories ranged from 0% to 9.86% at baseline and from 0% to 5.41% at follow-up.

Measurement of concordance

The marginal size of the black areas in the agreement charts (**Fig. 1**) in relation to the total size of the respective rectangles shows the low concordance on the diagonals that could already be read from Table I. There was very little concordance for the rating criteria at the ends of the spectrum. Agreement between patients and dermatolo-

Table I. Cross-tabulation of self-reported health (SRH) and physician-reported health (PRH) values at baseline and after 2 years

	Baseline					2-year follow-up				
	1	2	3	4	5	1	2	3	4	5
1	0.64%	1.59%	3.26%	3.42%	1.19%	0%	0.64%	8.25%	21.78%	13.14%
2	1.75%	9.86%	8.19%	3.58%	0.4%	0.13%	5.41%	11.73%	7.09%	1.8%
3	4.13%	20.19%	5.8%	0.79%	0.24%	0.64%	9.54%	5.41%	1.16%	0.26%
4	6.76%	10.1%	2.7%	0.32%	0.08%	0.64%	5.54%	1.42%	0.52%	0.26%
5	9.3%	4.45%	1.11%	0.16%	0%	1.8%	1.93%	0.64%	0.13%	0.13%

1=clear, 2=almost clear, 3=moderate, 4=severe, 5=very severe.
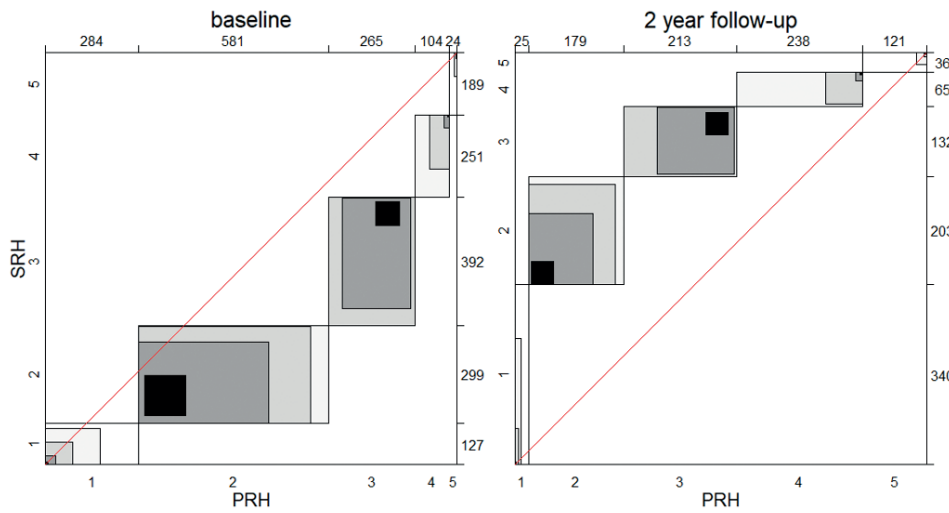
Fig. 1. Agreement chart of self-reported severity (SRH) and physician-reported severity (PRH) values at baseline and after 2 years (1="clear", 2="almost clear", 3="moderate", 4="severe", 5="very severe") (black = SRH and PRH identical, dark grey= SRH and PRH differ by one severity categorie, grey= SRH and PRH differ by two severity categories, light grey= SRH and PRH differ by three severity categories.

gists was found only in the categories "almost clear" (2) and "moderate" (3).

At baseline the rating mostly differs by only 1 category (e.g. a PRH rating of 2 and SRH rating of 3 result in deviation between the raters of 1. This accounts for a larger proportion of the dark-grey rectangle in the corresponding area). Concordance and slight deviations almost fill the entire field. Complete concordance was 16.62% at baseline and 11.47% at follow-up.

The discrepancies increased significantly in the marginal categories, indicated by a larger proportion of lighter grey areas. Extreme cases were found in which patients had classified their hand eczema as "clear" (1), whereas dermatologists had classified it as "very severe" (5) and vice versa. At baseline this proportion is 10.49%; at follow-up it is 15.07%.

Divergent trends for the paths of the rectangles were evident for both time-points. While the assessment of severity by the patients at baseline was relatively homogeneously distributed over the 5 categories (between 10% and 30% of the total), 66% of assessments

by the dermatologists were concentrated in the middle categories 2 and 3. A shift therefore occurred towards the coordinate (0, n) below the diagonal. At follow-up, the dermatologists' assessments were distributed relatively homogeneously (between 15.6% and 30% of the total) except for the category "clear", while 70% of the patients rated their hand eczema as "clear" or "almost clear".

Overall, patients rated their skin condition at study entry as more severe than did their treating dermatologists. After completion of the study period the remaining patients subjectively experienced significantly less severe hand eczema than was assessed by their dermatologists. Bangdiwala's weighted B statistic yielded a value of $\hat{B}_\omega = 0.573$ at basline and $\hat{B}_\omega = 0.498$ at follow up.

In contrast to Cohen's kappa, Bangdiwala's B has empirically derived interpretations of its values, which follow the categories of kappa as established by Landis & Koch (34). With values above 0.49, there was a low substantial agreement between the ratings at both points in time.
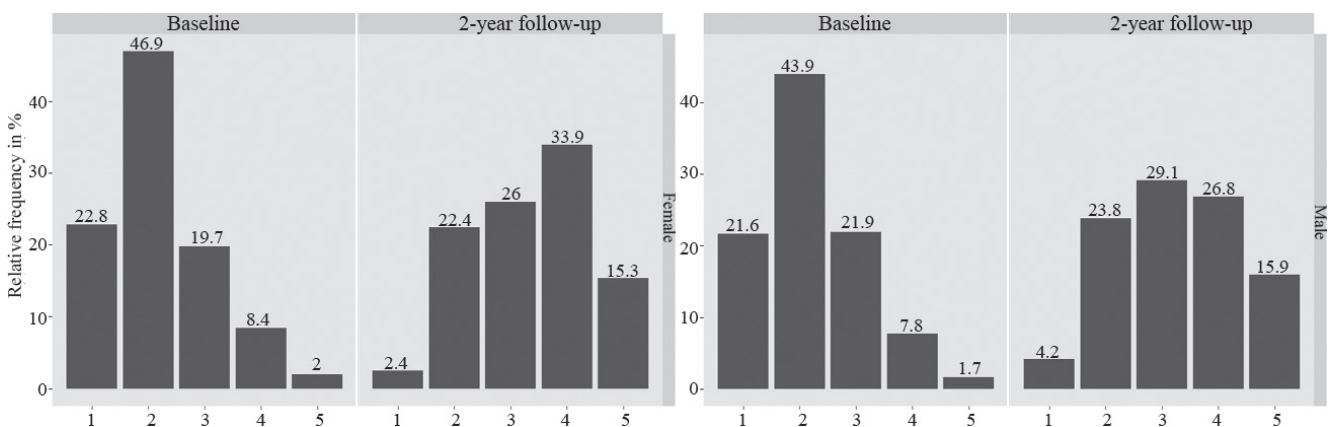
Fig. 2. Relative frequency of self-reported health (SRH) values by sex at baseline and after 2 years. (1="clear", 2="almost clear", 3="moderate", 4="severe", 5="very severe"). Regarding the chosen severity, it could be seen that women assessed their hand eczema as "severe" or "very severe" more often than men after 2 years and only half as often as "clear". These differences could not be shown for baseline.

Socio-demographic influences

At baseline, women showed a lower concordance with their dermatologists regarding assessment of CHE than did men (women: $\hat{B}_\omega = 0.562$, $n = 688$; men: $\hat{B}_\omega = 0.571$, $n = 593$). However, the differences between the values for both sexes were not striking and the deviations did not affect the qualitative assessment of the agreement. After 24 months, however, men fell below the threshold of 0.49 to only moderate agreement (34) compared with women (women: $\hat{B}_\omega = 0.494$, $n = 419$; men: $\hat{B}_\omega = 0.468$, $n = 357$). Overall, however, no substantial correlation of sex as the sole criterion and rating concordance was evident (**Fig. 2**).

It could be seen that the concordance at baseline decreased steadily with increasing age. Younger people (age range 18–30 years) agreed to a greater extent with the assessment by the treating dermatologist than did older people. In addition, women showed consistently lower agreement with the assessment by the treating dermatologist than did men. The difference, however, was less clear at 24-month follow-up. Women up to 30 years of age showed a much higher agreement with the assessment by the treating dermatologist than did men of the same age (women: $\hat{B}_\omega = 0.643$, $n = 64$; men: $\hat{B}_\omega = 0.534$, $n = 45$). The same was true for women and men aged 66-75 years (women: $\hat{B}_\omega = 0.408$, $n = 29$; men: $\hat{B}_\omega = 0.268$, $n = 22$) (**Fig. 3**).

The observed general tendency for patients to rate their CHE as significantly more severe than did the treating dermatologist at the time of the initial survey was consistent across both sexes and all age cohorts, as was the inversion of the relationship at 2-year follow-up. Dermatologists' and patients' ratings at follow-up were also more likely to differ from each other by more than 1 severity rating category.

DISCUSSION

A number of studies referring to patient's general health have investigated concordance between dermatologists and patients assessments (16–19, 21). The concordance rates described range from 37.8% (21) to 59.6% (18), which is significantly higher compared with the results of the current study. Apart from one study (19), the study populations were always older individuals (58+ years).

Three other studies included patients with atopic dermatitis (AD) (35–37), one included patients with psoriasis (38). One compared the assessment of the current severity of AD on a 3-point scale in 687 dermatologist-patient pairs. Here, 11.2% of the patients evaluated the disease as more severe, and 20.2% as less severe compared with their dermatologists' evaluation. Concordance was 68.6% (37). Two other AD studies compared concordance by other means, but also reported higher severity ratings by patients (35, 36). Another study compared the self-assessments of 502 patients with psoriasis using the PGA score, and found 13% overestimation and 26% underestimation of severity compared with the dermatologists' assessments. Complete concordance was 61% (38). The agreement values in all four studies were higher than those described here.

A self-assessed "overestimation" of the patient's health and an underestimation of the severity of CHE compared to the dermatologists assessment could be shown in the present analysis. This is in line with the findings of an other studie that found twice as high proportions of overestimation than underestimation (9). The current study even found a ratio close to 3:1. This may be an effect due to participation in this registry study and the associated expectations among patients of treatment and success. Inclusion in a registry study may itself serve as
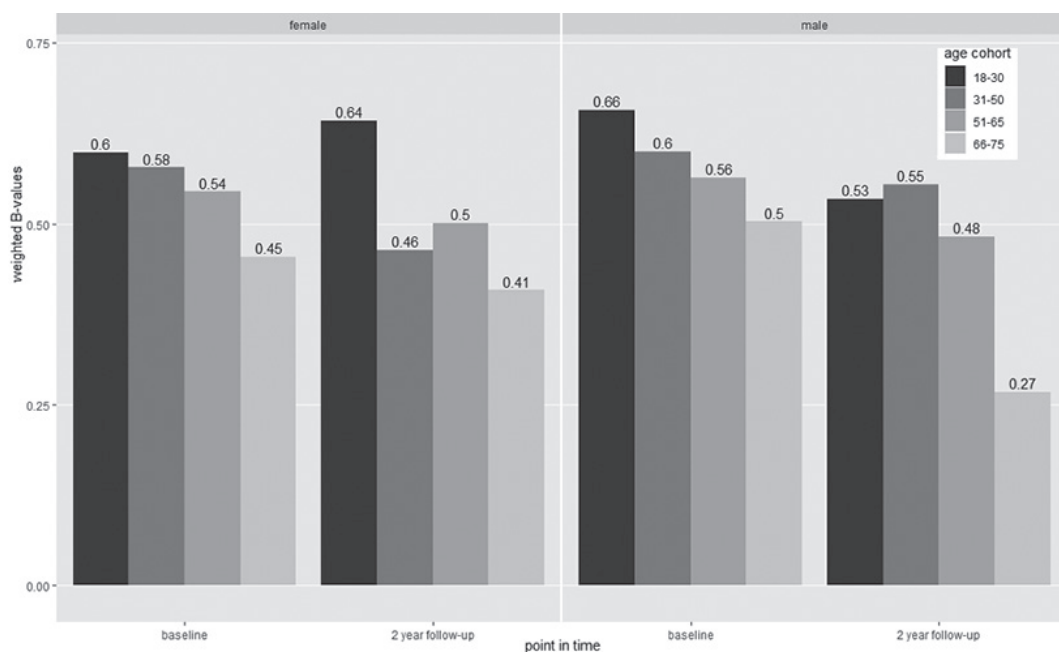
Fig. 3. Weighted B-values for concordance by sex and age cohort at baseline and after 2 years (B > 0.49 = substantial agreement, B < 0.49 = moderate agreement).

a kind of "therapy": The CARPE-patients were regularly scheduled for study visits and received close evaluation of their skin disease and treatments. This may raise expectations of success similar to a placebo effect. These expectancy effects have been described in the field of psychotherapeutic interventions (39).

The sex differences in the assessment of severity of CHE in the current study were consistent with study results on general health and pain perception among men and women. Women reported significantly worse SRH scores than men (40), which may have its origin in social norms and gender-stereotypical behaviour in environments such as the workplace (41). The influence of the patient's sex on general pain perception showed that women exhibit higher sensitivity to pain (42, 43). This may also have a decisive influence on the assessment of the subjectively perceived severity of CHE. It can be hypothesized that women might perceive their CHE to be more painful than men and therefor rate it as more severe.

Study limitations and strengths

A limitation of the current analysis is that the assessment scales of patients and dermatologists were not congruent. While dermatologists were asked to divide their measurements into 5 fixed categories according to uniform criteria using the PGA, patients marked their assessment on a VAS. The transfer into 5 uniform categories was linear, according to purely metric criteria. It remains open whether this static subdivision is theoretically tenable or whether the perception of severity follows non-linear criteria leading to differently broad assessment categories. One published study shows evidence that the recent state of the skin disease could affect the patient's recollection of former states and therefore influence a VAS rating (44). Even though not explicitly asked to rate the severity of the condition compared with earlier states, patients will be likely to use them as reference for their answers. This might have influenced their answers. In accordance with the comment on another study (38), it should be emphasized that the patient's perspective often incorporates different, non-medical aspects into the assessment of severity or general health. This does not mean that patient-related outcomes are less valid, but that they often do not receive the same awareness as the dermatologists' perspective. Concordance analyses on health or disease state therefore inevitably suffer from validity problems from a measurement-theoretical perspective. The PGA rating focuses solely on the morphology of the CHE and is also subjective in nature, which can increase discordance even further. Also, different levels of experience among the dermatologists with CHE and the PGA rating could have influenced the results.

The strength of the current study is the extensive data basis of the analyses, which is almost twice as large as in comparable studies. Using follow-up data to analyse the changes in the patients' assessment of severity of their CHE gives credit to the development over time of CHE under treatment. Using a concordance measurement that takes the special structure of the data into account and includes graphical supplementation, underlines the reliability of the findings and is superior to the commonly used kappa statistic.

Conclusion

This study demonstrates differences in the assessment of CHE between patients (especially women) and dermatologists. In conclusion, dermatologists should focus in particular on "health optimists" (9), patients who assess their disease as less severe than it is clinically classified by dermatologists and who might show less adherence to treatment because they consider the need for therapy differently. At the same time, "health pessimists" may have the impression that their dermatologist does not take their disease perception seriously. It is recommended that dermatologists ask the patients about their subjective perception of the disease independently of severity and treatment, as this may improve therapy success.

REFERENCES

1. Berthold E, Weisshaar E. Therapie des Handekzems. Hautarzt 2019; 70: 790–796.
2. Carroll CL, Balkrishnan R, Feldman SR, Fleischer AB, Jr., Manuel JC. The burden of atopic dermatitis: impact on the patient, family, and society. Pediatr Dermatol 2005; 22: 192–199.
3. Kleyn CE, Barbarot S, Reed C, Losi S, von Arx LB, Robert C, et al. Burden of moderate to severe atopic dermatitis in adults from France, Italy, and the UK: patient-reported outcomes and treatment patterns. Dermatol Ther 2022; 12: 1947–1965.
4. Ofenloch RF, Weisshaar E. Epidemiologie des Handekzems in Deutschland. Ein Rückblick auf die letzten 10 Jahre Handekzemforschung in Deutschland. Hautarzt 2019; 70: 766–772.
5. Diepgen TL, Andersen KE, Chosidow O, Coenraads PJ, Elsner P, English J, et al. Guidelines for diagnosis, prevention and treatment of hand eczema – short version. J Dtsch Dermatol Ges 2015; 13: 77–85.
6. Tian M, Chen Y, Zhao R, Chen L, Chen X, Feng D, et al. Chronic disease knowledge and its determinants among chronically ill adults in rural areas of Shanxi Province in China: a cross-sectional study. BMC Public Health 2011; 11: 948.
7. Slama-Chaudhry A, Golay A. Patient education and self-management support for chronic disease: methodology for implementing patient-tailored therapeutic programmes. Public Health Panor 2019; 5.
8. Feng Q, Zhu H, Zhen Z, Gu D. Self-rated health, interviewer-rated health, and their predictive powers on mortality in old age. J Gerontol B Psychol Sci Soc Sci 2016; 71: 538–550.
9. Geest TA, Engberg M, Lauritzen T. Discordance between self-evaluated health and doctor-evaluated health in relation to general health promotion. Scand J Prim Health Care 2004; 22: 146–151.
10. Wuorela M, Lavonius S, Salminen M, Vahlberg T, Viitanen M, Viikari L. Self-rated health and objective health status as predictors of all-cause mortality among older people: a prospective study with a 5-, 10-, and 27-year follow-up. BMC Geriatr 2020; 20: 120.
11. Baron-Epel O. Self-reported health. In: Anderson NB, editor. Encyclopedia of health and behavior. New York: SAGE Publications; 2004: p. 630–634.
12. Nielsen AB, Siersma V, Hiort LC, Drivsholm T, Kreiner S, Hollnagel H. Self-rated general health among 40-year-old Danes and its association with all-cause mortality at 10-, 20-, and 29 years' follow-up. Scand J Public Health 2008; 36: 3–11.
13. Zhu H, Feng Q, Gu D. Self-rated health and interviewer-rated health: differentials in predictive power for mortality among subgroups of Chinese elders. Int J Popul Stud 2016; 2: 74–91.
14. Eriksson HG, von Celsing AS, Wahlstrom R, Janson L, Zander V, Wallman T. Sickness absence and self-reported health a population-based study of 43,600 individuals in central Sweden. BMC Public Health 2008; 8: 426.
15. Linn BS, Linn MW. Objective and self-assessed health in the old and very old. Soc Sci Med Med Psychol Med Sociol 1980; 14A: 311–315.
16. Borawski EA, Kinney JM, Kahana E. The meaning of older adults' health appraisals: congruence with health status and determinant of mortality. J Gerontol B Psychol Sci Soc Sci 1996; 51: S157–170.
17. Desalvo KB, Muntner P. Discordance between physician and patient self-rated health and all-cause mortality. Ochsner J 2011; 11: 232–240.
18. Idler EL, Hudson SV, Leventhal H. The meanings of self-ratings of health: a qualitative and quantitative approach. Res Aging 1999; 21: 458–476.
19. Kivinen P, Halonen P, Eronen M, Nissinen A. Self-rated health, physician-rated health and associated factors among elderly men: the Finnish cohorts of the Seven Countries Study. Age Ageing 1998; 27: 41–47.
20. Klein R, Scherer J, Meyer F. Die Krankheit aus der Sicht von Patient und Arzt - wie beurteilen beide die Situation?. ZFA 2003; 79: 248–253.
21. van Doorn C. A Qualitative approach to studying health optimism, realism, and pessimism. Res Aging 1999; 21: 440–457.
22. Apfelbacher C, Molin S, Weisshaar E, Bauer A, Elsner P, Mahler V, et al. Characteristics and provision of care in patients with chronic hand eczema: updated data from the CARPE registry. Acta Derm Venereol 2014; 94: 163–167.
23. Apfelbacher C, Weiss M, Molin S, Bauer A, Mahler V, Schmitt J, et al. Which factors are associated with the use of systemic antihistamines in patients with chronic hand eczema? Results from the CARPE registry. J Eur Acad Dermatol Venereol 2016; 30: 50–56.
24. Apfelbacher CJ, Akst W, Molin S, Schmitt J, Bauer A, Weisshaar E, et al. CARPE: a registry project of the German Dermatological Society (DDG) for the characterization and care of chronic hand eczema. J Dtsch Dermatol Ges 2011; 9: 682–688.
25. Apfelbacher CJ, Ofenloch RF, Weisshaar E, Molin S, Bauer A, Mahler V, et al. Chronic hand eczema in Germany: 5-year follow-up data from the CARPE registry. Contact Dermatitis 2019; 80: 45–53.
26. Ruppert L, Apfelbacher C, Molin S, Bauer A, Mahler V, Schmitt J, et al. Itching in patients with chronic hand eczema: data from the CARPE registry. Dermatology 2014; 229: 146–153.
27. Shankar V, Bangdiwala SI. Behavior of agreement measures in the presence of zero cells and biased marginal distributions. J Appl Stat 2008; 35: 445–464.
28. Bangdiwala SI. Graphical aids for visualizing and interpreting patterns in departures from agreement in ordinal categorical observer agreement data. J Biopharm Stat 2017; 27: 77–783.
29. Cicchetti DV, T A. A new procedure for assessing reliability of scoring EEG sleep recordings. Am J EEG Technol 1971; 11: 101–110.
30. Kassambara A. Weighted Kappa in R: For Two Ordinal Variables. 2018 accessed 21.02.2023, https://www.datanovia.com/en/lessons/weighted-kappa-in-r-for-two-ordinal-variables/.
31. Eubanks D. (Re)visualizing rater agreement: beyond single-parameter measures. J Writ Anal 2017; 1: 276–310.
32. Bangdiwala SI, Shankar V. The agreement chart. BMC Med Res Methodol 2013; 13.
33. Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the Tidyverse. J Open Source Softw 2019; 4: 1686.
34. Meyer D, Zeileis A, Hornik K, Gerber F, Friendly M. vcd: Visualizing categorical data: R package version 1.4-8; 2020.
35. Munoz SR, I BS. Interpretation of kappa and B statistics measures of agreement. J Appl Stat 1997; 24: 105–112.
36. Cvetkovski RS, Jensen H, Olsen J, Johansen JD, Agner T. Relation between patients' and physicians' severity assessment of occupational hand eczema. Br J Dermatol 2005; 153: 596–600.
37. van Coevorden AM, van Sonderen E, Bouma J, Coenraads PJ. Assessment of severity of hand eczema: discrepancies between patient- and physician-rated scores. Br J Dermatol 2006; 155: 1217–1222.
38. Wei W, Anderson P, Gadkari A, Blackburn S, Moon R, Piercy J, et al. Discordance between physician- and patient-reported disease severity in adults with atopic dermatitis: a US cross-sectional survey. Am J Clin Dermatol 2017; 18: 825–835.
39. Griffiths C, Naldi L, Augustin M, Romiti R, Guevara-Sanginés E, Howe T, et al. Discordance between patients with psoriasis and their physicians in assessing disease severity, psoriasis-related symptoms and disease control: results from a large observational physician and patient survey. J Am Acad Dermatol 2017; 76: AB102.
40. Rief W, Glombiewski JA. Erwartungsfokussierte Psychotherapeutische Interventionen (EFPI). Verhaltenstherapie 2016; 26: 47–54.
41. Boerma T, Hosseinpoor AR, Verdes E, Chatterji S. A global assessment of the gender gap in self-reported health with survey data from 59 countries. BMC Public Health 2016; 16: 675.
42. Caroli E, Weber-Baghdiguian L. Self-reported health and gender: the role of social norms. Soc Sci Med 2016; 153: 220–229.
43. Bartley EJ, Fillingim RB. Sex differences in pain: a brief review of clinical and experimental findings. Br J Anaesth 2013; 111: 52–58.
44. Fillingim RB, King CD, Ribeiro-Dasilva MC, Rahim-Williams B, Riley JL, 3rd. Sex, gender, and pain: a review of recent clinical and experimental findings. J Pain 2009; 10: 447–485.
45. Mollerup A, Johansen JD. Response shift in severity assessment of hand eczema with visual analogue scales. Contact Dermatitis 2015; 72: 178–183.