Eur J Dermatol 2009; 19 (3): 224-31

Johannes GASSMUELLER
Elisabeth ROWOLD
Thomas FRASE
Betsy HUGHES-FORMELLA

bioskin GmbH, Burchardstraße 17,
D-20095 Hamburg, Germany

Reprints: B. Hughes-Formella
<betsy.hughes@bioskin.de>

Validation of TrichoScan® technology as a fully-automated tool for evaluation of hair growth parameters

There is a need for a simple and reliable tool to evaluate hair loss and treatment effects in patients suffering from alopecia. In 2001 TrichoScan® was introduced as a fully-automated method for the measurement of biological parameters of hair growth such as density, diameter and growth rate. However, the conventional phototrichogram method with manual marking of hairs on images is still performed and, although no real independent side-by-side comparison is available, the manual method is sometimes defined as the most precise method of measurement. The aim of this study was validation of the Trichoscan® method by comparative assessment of TrichoScan® analysis and manual marking of hairs. Digital images were taken from 10 patients with androgenetic alopecia (AGA) and validity and reliability of both methods were assessed. This study showed an excellent correlation of TrichoScan® and manual marking of hairs. Considerable variability was noted in the results from manually evaluated images (range 2.71%-12.95%), compared to none in TrichoScan® analyzed images. Results with TrichoScan® were obtained more quickly and were more reproducible with a smaller margin of operator error. The consistency in the Trichoscan® data allows statistically significant results to be obtained with a smaller sample size.

Article accepted on 24/12/2008

Key words: androgenetic alopecia, TrichoScan, hair, measurement

In 2001 TrichoScan® was introduced as a fully-automated method for the measurement of biological parameters of hair growth such as hair density or hair diameter [1]. The method combines epiluminescence microscopy (ELM) with automatic digital image analysis for the measurement of human hair [2]. Whereas in the past the dermatologist or technician literally had to measure the length of the hair fibers and diameters with a ruler, image analysis programs now allow for easy analysis of the marked hairs. However, the manual identification of the hairs is a tedious process prone to human error, even though manual identification of hairs is sometimes defined as the most precise method of measurement [3]. This might be the case when only the number of hairs in a very small scalp area is counted, but in large areas it is difficult for even well-trained technicians to make accurate manual counts of the total, terminal, and vellus hairs as well as hair thickness and hair growth rate. This must result in variable results when the same image is counted two or more times. Since at present no real independent side-by-side comparison is available. Therefore, the aim of this study was the validation of the TrichoScan® method by comparative assessment of hair growth parameters using TrichoScan® software *versus* manual identification of hairs prior to the final assessment of hair parameters.

Abbreviations:

GCP	Good Clinical Practice
CRO	Contract Research Organization
AGA	Androgenetic or pattern Alopecia

Materials and methods

Study participants

Digital images for TrichoScan® or conventional visual analysis were taken from 10 patients aged 18 years or older with AGA, Norwood-Hamilton grade III-IV/Ludwig grade 1 or 2. The participants were selected from the volunteer panel at bioskin GmbH, Hamburg. All patients included in this study had also taken part in a previous hair growth study at bioskin and already had a suitable measurement area marked with a tattoo. There were no other pigmented lesions in the treatment area. The recommendations of the Helsinki Declaration and the ICH GCP guidelines were followed. Written informed consent was obtained before inclusion in the study.

Tracking of target area and hair clipping

On day 1 the measurement area on the anterior border of the vertex balding spot was identified. The area was clipped evenly (Moser, TrichoScan Edition) and short clipped hair was removed by pressing an adhesive strip onto the shaved area three times. The quality was checked

with a magnifying glass. Afterwards a digital image was taken for documentation of the time. Digital images were stored in an image database (Image DB).

Hair dye and TrichoScan® image on day 3

On day 3 (48 ± 2 hours after hair clipping) the clipped hairs within the target area were dyed (Goldwell topchic, black 2N, Darmstadt, Germany with Rondo 6% Crème-Oxyd, Coiffeur, Cologne, Germany). After 12 minutes the colored area was thoroughly cleaned with an alcoholic solution (Kodan® Spray, Schülke & Mayr, Vienna, Austria) and digital images were taken using a digital ELM system. Three separate images (B1, B2, B3) were taken by 3 investigators (U1, U2, U3).

Evaluability criteria – Images for TrichoScan® analysis

To be admissible the following requirements had to be fulfilled for all images: All hairs were uniformly dyed, all hairs were evenly clipped, no remnants of hair dye were present, no air bubbles were present around the hairs, the image was bright and sharp, no hairs from outside the measurement area crossed the field, and all hairs were straight. None of the images which were taken had to be excluded from analysis.

TrichoScan® analysis and generation of data base

All images were analyzed using TrichoScan® Research Edition 3.0 and results were imported into Excel. Data obtained by manual evaluation were extracted with special software into a tab delimited text file which was also imported into an Excel® data sheet. The statistical analysis was performed at bioskin.

Conventional image analysis by hand (manual evaluation)

Three CDs of the images were produced for manual analysis by three independent evaluators. These CDs contained the same images which were analyzed by TrichoScan®, but they were additionally embedded in a software program ("hair measure tool" provided by Datinf GmbH, Tübingen). This software contained all 90 images in random order. Randomization was done by Datinf GmbH, Tübingen, Germany. The images were numbered 1, 2, (x), – 90. No information was given about who took the image and from which subject the image was taken. It was not possible to delete images from or to add images to this CD. Each evaluator used a computer mouse to outline the perimeters of each hair fiber on each image. He/she had to click on every hair where the hair left the scalp skin, then follow the hair and release the mouse button at the end of the hair tip. All such marked hairs then appeared in yellow in the software program *(figure 1)*. The thickness of the yellow line (the hair) was adjusted to the actual hair thickness with the scroll wheel on the computer mouse. When the yellow line had the same thickness as the actual underlying hair, the correct thickness of this hair was determined. Hair density (number per unit area on the image) and hair thickness (hair diameter) were recorded automatically by the software. Hairs starting from outside the target area which had the hair tips inside the target area were not counted. Hairs which started inside the target area but left it were counted, however these were not used for the analysis of hair growth as the complete hair shaft was not in the target area. The evaluator did not have access to any of the calculated results such as hair density and thickness, therefore he/she was unable to compare different analyses. After each analysis the manual evaluator had to click "finish" and thereafter this image was no longer available for counting. No information was given about the results of the computerized TrichoScan® images.

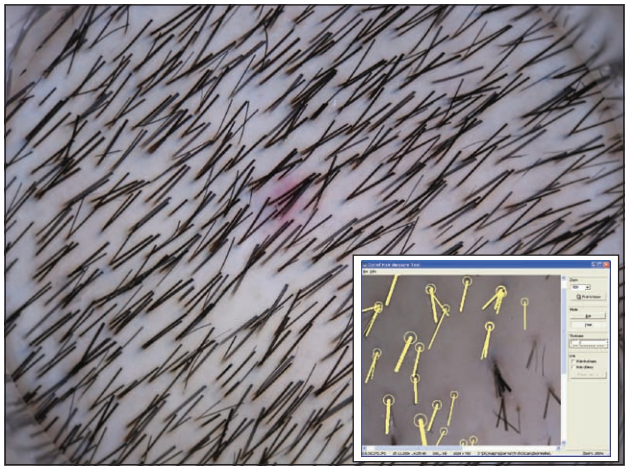

Figure 1. Dermatoscopic image from the scalp. Insert shows a screenshot of the software for manual hair evaluation. Yellow: marked hairs; no color: hairs to be analyzed.

Data analysis and statistics

Study objectives

Validity

To prove the validity of TrichoScan® it was necessary to show a strong correlation with manual evaluation. All analyses were done with images as unit of observation and with patients as unit of observation. The former ensured that all images were analyzed without respect to multiple images of one subject. However, since in practice not the images but the patients are the relevant reference, it was also necessary to do the analysis for the values averaged for patients.

Reliabilty

To evaluate the reliability, three different sources of variation had to be taken into account: the patient, the investigator taking the image, and the evaluator doing the manual evaluation or the TrichoScan® software. Analysis of variability was done separately for the manual evaluation of hair parameters and TrichoScan®.

Number of evaluations

Images were taken from 10 patients. For each patient three investigators took three images each on day 3, for a total of 90 different images. Each image was analyzed manually by three different evaluators (270 manually analyzed images) as well as three times with TrichoScan®. In addition, the first image for each patient was made by Investigator 1 two more times. This means that those images were analyzed manually three times. The same was done with TrichoScan®. A total of 660 evaluations were performed, 50% by hand, 50% by TrichoScan® *(tables 1 and 2)*. The mean hair den-

Table 1. Overview of images, evaluation/analysis of images and additional repetitions: Number of different images was $90 = 10 \times 3 \times 3$ (product of no. of levels of SN, Inv and Img); Number of evaluations without "Additional" was $90 \times 2 \times 3 = 540$ (product of different images and no. of levels of Met and EvalNo); Number of "Additional" evaluations was $10 \times 2 \times 3 \times 2 = 120$ (product of different images and no. of levels of Met, EvalNo and repetitions (2))

	Description	Variable	Levels	No. Levels
Images	Patient number	SN	1.10	10
	Investigator defined as person who took the images	Inv	1,2,3	3
	Repetition of image from Investigator	Img	1,2,3	3
Evaluation/Analysis of images	Method, TrichoScan® (TS) or manual analysis (Hand) of image	Met	TS, Hand	2
	No. of evaluation (three TrichoScan® analyses and three different evaluators for manual hair counting	EvalNo	1,2,3	3
Additional	Repetition of evaluation – only for Met Hand, Img 1 and Inv 1, *i.e.* two additional repeated analyses	RepNo	1,2,3	3

sity was 199/cm^2 and 223/cm^2, respectively. As we analyzed an area of 1.42 cm^2, this adds up to 282 and 316 hairs/image. This number multiplied by the number of evaluations amounts to a total of 198,018 analyzed hairs.

Analyzed variables

The following variables were evaluated manually and by TrichoScan® for each image: TotalDens (total hair density, n/cm^2); TerminalDens (density of hairs thicker than 40 μm, n/cm^2); CumThickTotal (cumulative thickness of all hairs, mm/cm^2); CumThickTerm (cumulative thickness of all terminal hairs, mm/cm^2); MeanThickTotal (mean thickness of all hairs, μm); MeanThickTerm (mean thickness of all terminal hairs, μm); GrowthRateTotal (mean length of all hairs, mm/day); GrowthRateTerm (mean length of all terminal hairs, mm/day); CumGrowthRateTotal (sum length of all hairs, mm/day); and CumGrowthRateTerm (sum length of all terminal hairs, mm/day).

Table 2. Overview of data resulting from a) different images and b) additional evaluation of image 1 by Investigator 1 to investigate the effect of repeated analyses of identical images

a) All different images (90)			
Met	**Eval No.**	**Rep No.**	**No. of evaluations**
TrichoScan	1	1	90
TrichoScan	2	1	90
TrichoScan	3	1	90
Hand	1	1	90
Hand	2	1	90
Hand	3	1	90
Sum			540
b) Additional repeated analysis of images of Inv 1, Img 1 (10)			
Met	**Eval No.**	**Rep No.**	**No. of evaluations**
TrichoScan	1	2	10
TrichoScan	1	3	10
TrichoScan	2	2	10
TrichoScan	2	3	10
TrichoScan	3	2	10
TrichoScan	3	3	10
Hand	1	2	10
Hand	1	3	10
Hand	2	2	10
Hand	2	3	10
Hand	3	2	10
Hand	3	3	10
Sum			120

Statistical methods

Validity

Descriptive statistics were performed, including differences between the two methods of evaluation (mean and standard deviation). In addition, a paired t-test (two-sided) was performed for the differences, and correlation coefficients (Pearson) were calculated. The analyses were done with images as the unit of observation and with patients as the unit of observation (*i.e.* the mean value calculated for each patient). Data from the repeated evaluations of image 1 of Investigator 1 were averaged before performing statistical tests.

Reliabilty

The analysis of the variation attributable to the investigator and evaluator was done according to Bland and Altman [4]. Briefly, the calculated variances were the variance between subjects (σ^2_b), observers (σ^2_o), different observers for different subjects (σ^2_h) and variance of observations by one observer for one subject (σ^2_w). A two-way analysis of variance (ANOVA) with the factors subject and observer and the interaction of subjects and observers was calculated. The intra-observer variability was σ^2_w, the inter-observer variability was $\sigma^2_o + \sigma^2_h + \sigma^2_w$. The repeatability by the same observer was calculated as $2.83 \times \sigma_w$ and the reproducibility when different observers were used was calculated as $2.83 \times$ square root $(\sigma^2_o + \sigma^2_h + \sigma^2_w)$. Lower values reflect high repeatability and reproducibility, higher values low repeatability and reproducibility. The repeatability (and reproducibility) was an estimate of the maximum difference (the limit within which 95% of differences will lie) which can be obtained between two measurements made at random on the same subject. The intra-class correlation (ICC) coefficient for a single observer was calculated as $\sigma^2_b/(\sigma^2_b + \sigma^2_w)$ and the ICC for different observers was calculated as σ^2_b/σ^2.

Intra- and inter-investigator variability and reliability. The first image of Investigator 1 was evaluated three times by both methods and the mean of the three repeated evaluations was calculated. The analysis was carried out for each evaluation method: Hand 1, Hand 2, Hand 3, TS 1, TS 2, TS 3.

Inter-evaluator variability. The first image of Investigator 1 was evaluated three times and the mean of the three

repeated evaluations was calculated. The mean from all three images taken by the investigator for one subject was used in the ANOVA. Since data from three different investigators were available, this analysis was performed three times using the data from each investigator. In addition the analyses were performed using the mean value of the investigators.

Intra- and inter-evaluator variability and reliability. A data subset was created using the first image from Investigator 1 with the three repetitive evaluations by the three evaluators. These data were used to calculate the intra- and inter-evaluator variability and reliability.

Overall (investigator and evaluator) variability and reliability. The overall variability and reliability were derived from the investigator and evaluator variability and reliability.

Variation coefficients. Variation coefficients were calculated from the intra-evaluator variability and the corresponding mean values. The results are given as percentages.

Results

Validity – Comparison of manual evaluation and TrichoScan®

There was a very strong correlation between the manually evaluated hair parameters and TrichoScan®, as evidenced by the high Pearson correlation coefficients. All correlations with patients as unit of observation were greater than 0.89 (0.85 with images as unit of observation). All correlation coefficients were highly significant ($p < 0.001$). The results are listed in *table 3*.

Reliability

Intra- and inter-investigator variability and reliability

A general impression can be gained from the results illustrated in *figure 2* for the plotted variables of total hair density and cumulative hair thickness. The greatest variance was associated with the patients. Therefore, the inter- and intra-investigator correlation coefficients are all in a high range. Furthermore, the variability for each patient was considerably higher for the individual manual evaluators than for TrichoScan®, resulting in lower repeatability and reproducibility, as well as lower intra-class correlation coefficients (ICC) for a single and for different investigators. The values are listed in *table 4*. The repeatability for the total hair density was approximately three to five times higher for TrichoScan® than for the three manual evaluators. Similar results were obtained for the other variables. In *table 4* the results are listed and it can be seen that in general TrichoScan® showed the highest intra-class correlation coefficients and the greatest repeatability and reproducibility. There was one exception: The "mean length of all hairs divided by time difference from shaving (mm/day)" was similar for TrichoScan® and the manual evaluators.

Intra- and inter-evaluator variability and reliability

Values for inter- and intra-evaluator variability can be found in *table 5*. As expected for a fully-automated sys-

Table 3. Comparison of manually (Hand) and TrichoScan® (TS) evaluated hair parameters: For each parameter the mean, the difference, and the relative difference in % are given for Hand and TS. These values are identical for the analysis with images as unit of observation and with patients as unit of observation. The correlation coefficient (Pearson) and the *p* value for the paired t-test between the two methods are given for the analysis with images as unit of observation and with patients as unit of observation

Data					Images (N = 90)		Patients (N = 10)	
Parameters	**Mean TS**	**Mean Hand**	**Diff TS-Hand**	**Rel. Diff TS-Hand**	**Correlation**	**p (Diff)**	**Correlation**	**p (Diff)**
Total hair density (n/cm^2)	199.2	223.8	– 24.6	– 11.6%	0.967***	< .0001	0.976***	0.001
Density of hair thicker as 40 µm (n/cm^2)	134.0	142.4	– 7.4	– 5.3%	0.956***	< .0001	0.980***	0.085
Cumulative thickness of all hairs (mm/cm^2)	14.95	14.71	0.24	1.6%	0.982***	0.003	0.996***	0.057
Cumulative thickness of all terminal hairs (mm/cm^2)	11.91	11.42	0.49	4.2%	0.943***	0.002	0.971***	0.200
Mean thickness of all hairs (µm)	54	48	6	12.3%	0.894***	< .0001	0.940***	< .0001
Mean thickness of all terminal hairs (µm)	62	56	6	10.5%	0.853***	< .0001	0.894***	< .0001
Mean length of all hairs divided by time difference from shaving (mm/day)	0.448	0.444	0.003	0.8%	0.978***	0.093	0.989***	0.510
Mean length of all terminal hairs divided by time difference from shaving (mm/day)	0.489	0.473	0.017	3.5%	0.980***	< .0001	0.992***	0.002
Sum length of all hairs divided by time difference from shaving (mm/day)	120.1	129.6	– 9.5	– 7.6%	0.978***	< .0001	0.983***	0.008
Sum length of all terminal hairs divided by time difference from shaving (mm/day)	90.19	89.87	0.32	0.4%	0.986***	0.73	0.996***	0.887

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

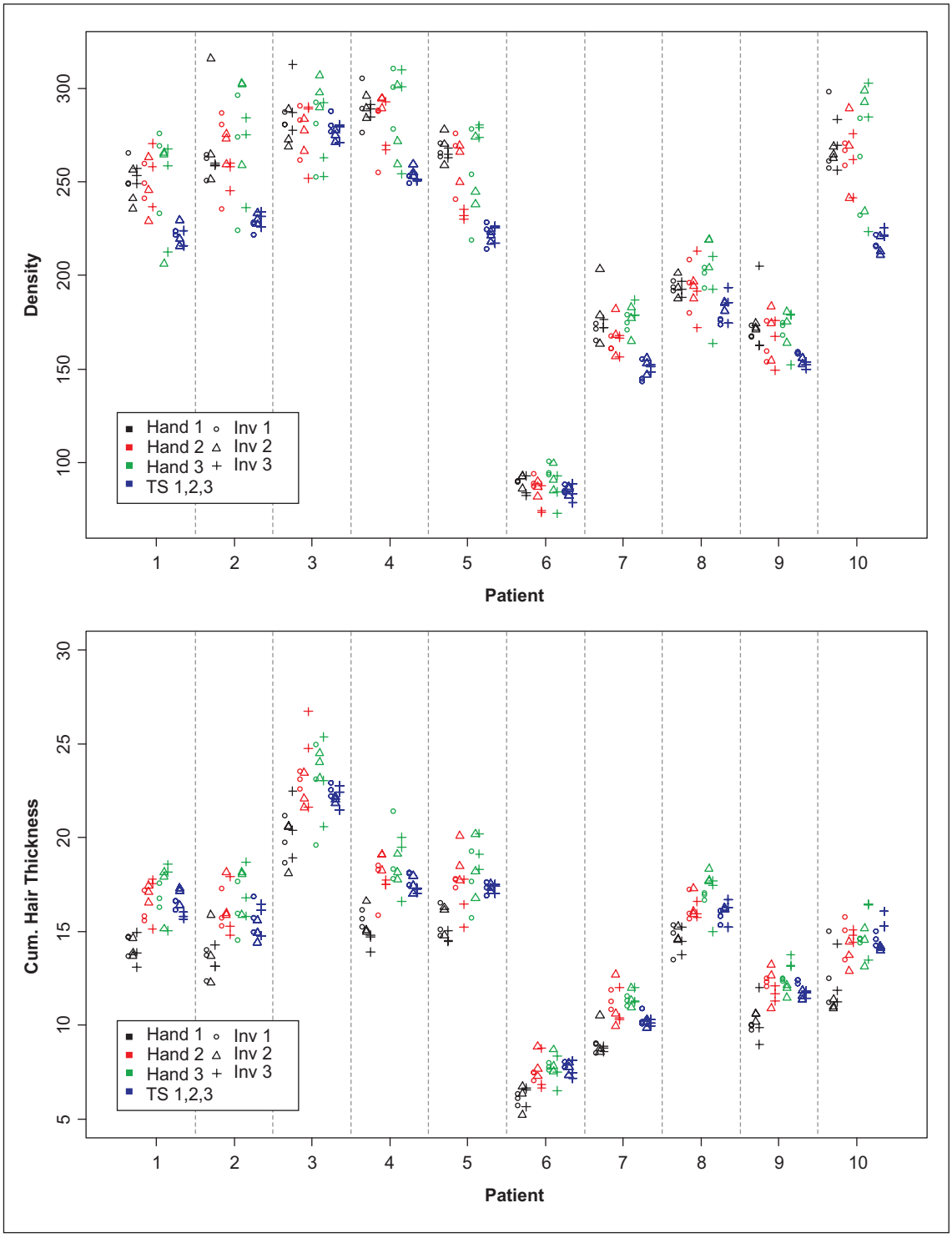

Figure 2. Hair density and cumulative hair thickness. For each method and evaluation (Hand 1-3 and TrichoScan® 1-3) and for each investigator the measured hair densities are shown using different symbols – color corresponds to the method, shapes to different investigators. Repeated symbols for one patient are due to the three repetitions by each investigator, *e.g.* the three black circles for patient 2 represent the three images taken by Investigator 1. The results of TrichoScan® evaluations 1 to 3 were identical and are shown as method TS 1,2,3. The results of the repeated evaluation of image 1 of Investigator 1 were averaged.

tem, there were no differences for TrichoScan® on repeated evaluations. Therefore the repeatability and reproducibility were always equal to 0 and the corresponding intra-class correlation coefficient was always equal to 1. This was not the case for the manual evaluation. In addition, the variation coefficients for intra-evaluator variability were calculated. Mean data variability in hand evaluated images for terminal hair thickness

Table 4. Results of all parameters for the inter- and intra-investigator variability for the three manual evaluations Hand 1, Hand 2 and Hand 3 and the three TrichoScan® evaluations which are identical and therefore listed only once as TS 1, 2, 3. For each parameter and evaluation the repeatability (maximum of the difference between two measurements on the same patient by the same investigator), the reproducibility (maximum of the difference between two measurements on the same patient by different investigators) and the corresponding intra-class correlation coefficients (ICC) are given

Parameter	Method	Repeatability by same investigator	Reproducibility by different investigators	ICC for a single investigator	ICC for different investigators
Total hair density (n/cm^2)	Hand 1	33.2	33.2	0.969	0.969
	Hand 2	39.3	40.5	0.956	0.953
	Hand 3	61.0	61.0	0.897	0.897
	TS 1, 2, 3	11.9	13.2	0.994	0.993
Density of hair thicker as 40 µm (n/cm^2)	Hand 1	39.1	39.2	0.918	0.918
	Hand 2	42.7	42.7	0.916	0.916
	Hand 3	45.5	45.8	0.907	0.906
	TS 1, 2, 3	15.7	18.8	0.981	0.972
Cumulative thickness of all hairs (mm/cm^2)	Hand 1	2.77	2.77	0.940	0.940
	Hand 2	2.95	3.00	0.944	0.943
	Hand 3	3.64	3.68	0.919	0.917
	TS 1, 2, 3	1.21	1.48	0.989	0.984
Cumulative thickness of all terminal hairs (mm/cm^2)	Hand 1	3.16	3.16	0.926	0.925
	Hand 2	4.36	4.36	0.889	0.889
	Hand 3	4.18	4.27	0.897	0.894
	TS 1, 2, 3	1.57	1.95	0.977	0.965
Mean thickness of all hairs (µm)	Hand 1	0.0060	0.0062	0.914	0.909
	Hand 2	0.0106	0.0107	0.810	0.810
	Hand 3	0.0080	0.0087	0.869	0.847
	TS 1, 2, 3	0.0030	0.0040	0.968	0.944
Mean thickness of all terminal hairs (µm)	Hand 1	0.0050	0.0050	0.882	0.882
	Hand 2	0.0077	0.0077	0.811	0.811
	Hand 3	0.0045	0.0052	0.915	0.890
	TS 1, 2, 3	0.0029	0.0038	0.962	0.935
Mean length of all terminal hairs divided by time difference from shaving (mm/day)	Hand 1	0.0415	0.0459	0.975	0.969
	Hand 2	0.0412	0.0426	0.978	0.976
	Hand 3	0.0529	0.0544	0.962	0.960
	TS 1, 2, 3	0.0309	0.0332	0.985	0.983
Mean length of all hairs divided by time difference from shaving (mm/day)	Hand 1	0.0219	0.0258	0.992	0.990
	Hand 2	0.0305	0.0351	0.986	0.982
	Hand 3	0.0439	0.0467	0.971	0.968
	TS 1, 2, 3	0.0287	0.0318	0.984	0.981
Sum length of all hairs divided by time difference from shaving (mm/day)	Hand 1	16.2	16.2	0.985	0.985
	Hand 2	26.4	27.0	0.963	0.961
	Hand 3	40.5	40.5	0.914	0.914
	TS 1, 2, 3	8.5	9.5	0.996	0.995
Sum length of all terminal hairs divided by time difference from shaving (mm/day)	Hand 1	18.6	18.7	0.980	0.979
	Hand 2	24.8	24.9	0.967	0.967
	Hand 3	26.2	26.2	0.962	0.962
	TS 1, 2, 3	9.7	12.1	0.993	0.989

ranged up to 12.95%, whereas the TrichoScan® variability was zero *(table 6)*.

Discussion

The present study was designed to compare the variability of a semi-automated procedure with manual identification of hairs prior to analysis by hair growth software with a fully-automated procedure with recognition of hairs by the software. In this GCP-conform study the clinical trial situation was imitated where different investigators use the same equipment on different patients.

In this study there was a highly significant correlation between evaluation of hair parameters using manual identification of hairs and the fully-automated TrichoScan® method. This demonstrates that the TrichoScan® software, although working by statistics and mathematical approxi-

Table 5. Results of all parameters for the inter- and intra-evaluator variability for the manual evaluation (Hand) and the TrichoScan® (TS) evaluation. For each parameter and method the repeatability (maximum of the difference between two measurements on the same patient by the same evaluator), the reproducibility (maximum of the difference between two measurements on the same patient by different evaluators) and the corresponding intra-class correlation coefficients (ICC) are given

Parameter	Method	Repeatability by same evaluator	Reproducibility by different evaluators	ICC for a single evaluator	ICC for different evaluators
Total hair density (n/cm^2)	Hand TS	44.3 0	46.5 0	0.943 1	0.937 1
Density of hair thicker as 40 µm (n/cm^2)	Hand TS	45.7 0	95.2 0	0.899 1	0.671 1
Cumulative thickness of all hairs (mm/cm^2)	Hand TS	3.32 0	5.27 0	0.929 1	0.838 1
Cumulative thickness of all terminal hairs (mm/cm^2)	Hand TS	4.15 0	8.29 0	0.894 1	0.679 1
Mean thickness of all hairs (µm)	Hand TS	0.0087 0	0.0178 0	0.849 1	0.572 1
Mean thickness of all terminal hairs (µm)	Hand TS	0.0063 0	0.0095 0	0.832 1	0.687 1
Mean length of all terminal hairs divided by time difference from shaving (mm/day)	Hand TS	0.0451 0	0.0614 0	0.972 1	0.949 1
Mean length of all hairs divided by time difference from shaving (mm/day)	Hand TS	0.0340 0	0.0695 0	0.984 1	0.935 1
Sum length of all hairs divided by time difference from shaving (mm/day)	Hand TS	29.3 0	31.5 0	0.953 1	0.946 1
Sum length of all terminal hairs divided by time difference from shaving (mm/day)	Hand TS	23.8 0	52.0 0	0.969 1	0.866 1

mation, counts hairs and not artefacts. Nevertheless, there are some differences between the methods concerning absolute values. The strongest differences were seen in the mean thickness and total density of the hairs. TrichoScan® values for hair thickness were approximately 10% higher and density values approximately 10% lower than the values obtained by the manual evaluators. The differences in most other parameters were less than 5%. In particular, the critical parameter cumulative hair thickness was in good agreement with the manual evaluation. It is not surprising that TrichoScan® underestimates the total hair density: As a digital tool TrichoScan® relies on camera resolution. In this study resolution was 2 Megapixel, a resolution at which very thin (below 7 µm) hairs are not analyzed. In our opinion this is not a disadvantage as these hairs are not clinically relevant. Regarding the overestimation of hair thickness by TrichoScan®, it must be kept in mind that this is a systematic difference and occurs for every hair. Therefore, it has no impact on the TrichoScan® capability to measure changes in hair thickness. If needed, the thickness can be mathematically readjusted.

Table 6. Data variability of manual evaluation (Hand) and TrichoScan®. The data variability of one investigator is shown

Variable	Hand	TrichoScan®
Hair Count Total	7.07%	0.00%
Hair Count Terminal	11.41%	0.00%
Cumulative Thickness Total	8.09%	0.00%
Cumulative Thickness Terminal	12.95%	0.00%
Mean Thickness	6.53%	0.00%
Mean Thickness Terminal	3.99%	0.00%
Growth Rate Total	2.71%	0.00%
Growth Rate Terminal	3.38%	0.00%
Cumulative Growth Rate Total	8.12%	0.00%

Not surprisingly, considerable variability was noted for the manually marked images. Manual marking of hairs is tedious and time-consuming, making it near to impossible to repeatedly count hundreds of hairs without variation for density, length and thickness. The mean data variability for one evaluator who marked the same image three times ranged from 2.71-12.95%, depending on the parameter. Some evaluators showed more variability than others. The correlation between different evaluators was best for total hair density and parameters related to hair length and worst for the parameters related to hair thickness. There was no variability in repeated measures with TrichoScan®, the software delivered completely reliable results. Keeping in mind the definition of repeatability, the maximum of the difference between two measurements on the same patient, this is a tremendous difference between manual marking of hairs and TrichoScan®. In a clinical trial setting this would mean that a much larger sample size is required in the case of manual marking compared to fully-automated evaluation.

In order to analyze data reproducibility due to investigator variability, the situation was compared when one investigator captured a phototrichogram image of the same target area three times or three different investigators captured the same target area. The statistical data correspond to the intra- and inter-investigator error related to the taking of images. Results for all parameters measured using TrichoScan® software verified that images made by one

investigator were highly reliable, with a very high correlation (ICC $\geq$ 0.962) for all parameters. Likewise, images of the same target area made by different investigators showed similar reproducibility. Although the manual evaluations also produced very robust results, the data variability introduced by manual measurements led to a lower overall data reproducibility.

It must be kept in mind that the quality of the results with TrichoScan® or any other imaging tool depends on the quality of the images. Only technically correct images can deliver good results. Although in our experience the TrichoScan® procedure is relatively easy, some investigators may have dye remnants, unfocused images, or large air bubbles in the image. In clinical trials this may be overcome by adequate investigator training and strict quality control of incoming images. To achieve the quality of images used in this study approximately 3 hours of training were required per investigator or study nurse.

In summary, in this validation study we saw an excellent correlation of hair growth parameters analyzed using the fully-automated TrichoScan® method and manual marking of hairs prior to analysis. Considerable variability was seen in the results from manually identified hairs, compared to none in TrichoScan®-analyzed images. In a clinical trial, the wider margin of error and consequent data variability from manually evaluated images would necessitate a larger study sample size to overcome the effect of the variability on the statistical calculations. Therefore, the TrichoScan® technique is particularly suitable for clinical studies with treatment comparisons. That this is indeed the case can be seen in the results of a recent clinical trial evaluating changes of hair growth and thickness in 34 Minoxidil-treated men. Using the TrichoScan® method, treatment benefits could be seen 8 weeks after treatment initiation [5]. Moreover, TrichoScan® can be adopted to study the effect of drugs or laser treatment on hypertrichosis or hirsutism [6]. This, however, requires a different software algorithm which was outside the scope of this trial.

Last but not least, it should be kept in mind that even though phototrichograms or other hair analysis methods are important tools in the evaluation of hair loss treatments, in later phase clinical trials it is also important to assess other measures such as quality of life. ■

Acknowledgements. *bioskin GmbH as an independent CRO was contracted by the TrichoScan® manufacturer (Tricholog GmbH, Freiburg, Germany) to perform this validation trial. TrichoScan® raw data were provided by Datinf GmbH, Tübingen, Germany and raw data from the manual hair counting by Tricholog.*

References

1. Hoffmann R. TrichoScan: combining epiluminescence microscopy with digital image analysis for the measurement of hair growth in vivo. *Eur J Dermatol* 2001; 11: 362-8.

2. Hoffmann R. TrichoScan: a novel tool for the analysis of hair growth in vivo. *J Investig Dermatol Symp Proc* 2003; 8: 109-15.

3. Van Neste D, Trüeb RM. Critical study of hair growth analysis with computer-assisted methods. *JEADV* 2006; 20: 578-83.

4. Bland JM, Altman DG. Measurement error. *BMJ* 1996; 313: 744.

5. Gassmueller J, Hoffmann R, Webster A. Treatment of androgenetic alopecia using topical fulvestrant solution in males and postmenopausal females; results from two randomized, proof-of-concept studies. *British J Dermatol* 2008; 158: 109-15.

6. Hoffmann R. A 4-month, open-label study evaluating the efficacy of eflornithine 11.5% cream in the treatment of unwanted facial hair in women using TrichoScan. *Eur J Dermatol* 2008; 18: 65-70.