International Journal of Colorectal Disease (2018) 33:567–576
https://doi.org/10.1007/s00384-018-2988-8

ORIGINAL ARTICLE

The presence of occipital hair in the pilonidal sinus cavity—a triple approach to proof

Dietrich Doll[1,2,3] · F. Bosche[1] · A. Hauser[4] · P. Moersdorf[5] · I. Sinicina[6] · J. Grunwald[7] · F. Reckel[7] · M. M. Luedi[8]

Accepted: 13 February 2018 / Published online: 28 February 2018

Abstract

Purpose Hair in the pilonidal sinus is not growing within the sinus cavity, as hair follicles are not present there. Not few pilonidal patients do not have intergluteal hair, which is said to be the causative agent of folliculitis and pilonidal genesis. So, what is the real source of the hair forming the typical pilonidal hair nest?

Methods A trifold approach was used: First, axial hair strength testing of pilonidal hair and body hair harvested from head, lower back (glabella sacralis), and cranial third of intergluteal fold. Hair strength match was compared clinically. Second, comparative morphological examination by expert forensic biologist of hair from sinus and dorsal body hair. Third, statistical Bayesian classification of every single sinus hair based on its strength was done to determine the most probable region of origin.

Results Using clinical hair strength comparison, in 13/20 patients, head hair is the stiffest hair, followed by intergluteal hair. Only in 6/20 patients, this is the case with hair from the glabella sacralis. According to comparative morphological comparison, a minimum of 5 of 13 hair nests with possible hair allocation examined contain hair from the occiput. In 5/18 nests, hair could not be determined to a specific location though. Statistical classification with correction for multiple testing shows that 2 nests have hair samples that are at least 100 times more probable to originate from head or lower back than from intergluteal fold.

Conclusion We saw our null hypothesis that "hair in the sinus cavity is from the intergluteal region" rejected by each of three different approaches. There is strong evidence that occipital hair is present regularly in pilonidal sinus nests. We should start thinking of occipital hair as an important hair source for the development of the pilonidal hair nest.

Keywords Hair · Pilonidal sinus · Scalp · Occiput · Cut hair · Electron microscopy · Pathogenesis

✉ Dietrich Doll
ddoll@gmx.de

[1] Department of Procto-Surgery, St. Marienhospital Vechta, Academic Teaching Hospital of the Medizinische Hochschule Hannover, Vechta, Germany

[2] Chris Hani Baragwanath Academic Hospital, University of Witwatersrand Medical School, Johannesburg, Republic of South Africa

[3] PSD Research Group, Vechtaer Institut für Forschungsförderung e.V., Marienstr. 6-8, 49377 Vechta, Germany

[4] Bern University of Applied Sciences, Burgdorf, Switzerland

[5] Department of Trauma, Hand and Reconstructive Surgery, Saarland University, Kirrberger Str. 100, 66421 Homburg, Saar, Germany

[6] Institute of Forensic Medicine, Ludwig-Maximilians-University, Munich, Germany

[7] Bavarian State Criminal Police Office, Forensic Science Institute, Maillingerstr. 15, 80636 Munich, Germany

[8] Department of Anesthesiology and Pain Therapy, Bern University Hospital Inselspital, University of Bern, Bern, Switzerland

Introduction

An infectious focus in the "lower back" of a captain's only 21-year-old son has been resisting the treatment of Dr. Anderson, who was injecting silver nitrate and mercury chloride into the wound, but healed as soon as a nest of hair was removed from the cavity [3]. From 1847, when he wrote his letter to the Editor of the Boston Medical and Surgical Society, up to now the question of the source of hair has not been raised. Firstly, pilonidal sinus was thought to be due to a neural tube defect, as openings were always found in the midline [26]. This assumption was later challenged as a regular connection to the neural canal with consecutive meningitis could be denied [21]. Furthermore, skin appendices as sweat gland and hair follicles were not present within the sinus cavity [9]. And—although pilonidal sinus disease (PSD) may be present in utero—the main age segment to occur is with the beginning of puberty [4, 11]. Between 1941 and 1945, 78,924 soldiers of

the US army presented with pilonidal disease, and peak incidence was attributed due to hard and bumpy seats in the motorized warfare [7]. Although later proven wrong [11, 15], hygiene and trauma had been tagged to pilonidal sinus, and as soldiers had short hair, it was thought that hair was not an issue any more.

The war is over; we are not driving jeeps any more, and the pilonidal incidence is rising continuously [11, 14, 17]. Recently, the presence of multiple short hair fragments within the sinus nests has been presented [6]. Significantly stiffer hair has been found in PSD patients as compared to sex-, age-, and BMI-matched pairs [10].

If stiffer rootless hair fragments are present in the pilonidal cavity, where are they coming from? Of course, exotic sinus contents have been found—a 37-cm-long and most probable female hair in a male sinus [19], or a feather in the sinus of a 19-year-old girl [13]. But applies this to the more than 40,000 young adults with pilonidal sinus every year in Germany?

From barbers [27], sheep shearers [28], and dog shearers [20], we have learnt that occupational exposure to large quantities of cut hair fragments leads to increased hair injection at the rostral side of their bodies. Not a single report has been indicating a higher pilonidal incidence at the back side of these people. Thus, if we look for the source of pilonidal hair, it must be searched at the back side.

It has long been suspected that hair from head or back may be the source of pilonidal hair [2, 18, 23, 31]. Mounting evidence from both case reports [12] and current original research [6] suggest that hair from other origin than the intergluteal fold may be essential in the pathogenesis of PSD.

The purpose of this study was therefore to determine the source of pilonidal sinus hair found in the cavity during surgery.

Our null hypothesis was that "hair in the sinus cavity is from the intergluteal region."

Materials and methods

Ethics

The ethics committee of the medical association of Niedersachsen, Berliner Allee 20, 30175 Hannover, Germany (Prof. Dr. med. Andreas Creutzig, chair), fully and unanimously approved the study based on § 15 of the Niedersachsen Medical Association's professional code of conduct.

Hair harvesting

Pilonidal hair samples from inside of the pilonidal sinus cavity were collected intraoperatively from 20 individual PSD patients. All patients had received an asymmetric modified Limberg or modified Dufourmentel closure following in toto excision of the pilonidal sinus and its fistula system following toluidine blue injection. Explicit care was taken to ensure harvesting every single hair from the sinus cavity specimen and from the surrounding scar tissue if present. The sinus system was cut open taking utmost attention not to produce any cut hair by this procedure. All hair samples were separated, washed, dried, and stored in a plastic bag at room temperature hidden from sunlight. Hair was handled according to the established methods described previously [10].

Additionally, hair from every PSD patient was epilated from their dorsal part of the head exostosis "protuberantia occipitalis externa" (POE), the lumbar lower back region "glabella sacralis" (GS), and the cranial third of the intergluteal fold (IGF). From each region, six hairs were epilated (if present). Terminal hair was addressed as "hair," while miniscule vellus hair (round, white, and of 2–3 mm length) was not seen as such.

Excel and Graphpad Prism were used to collect and analyze data and to design the graphics. *T* tests were conducted to compare two groups of data, and significance was set at $p \geq 0.05$.

Hair strength testing

Axial (trajectory) hair strength testing was applied as published before [10].

Maximum hair strength results were compared between body regions and sinus hair. It was postulated that the region with the strongest hair would be able to produce the strongest hair present at time of surgery in the sinus cave. Hair comparison was done on clinical grounds, using the hair strength testing results.

Comparative morphological examination

Following hair strength testing, the identical sinus hairs and body hair samples were sent for further morphological comparison investigation to the Forensic Institute of the Bavarian State Criminal Police Office, Munich, Germany. Two scientists in charge of biological material and trace analysis (FR and JG) conducted their research independently from each other and without knowledge of our results.

First, hair samples of PSD hairs and comparative hairs from the relevant different body regions of each patient were selected. The gross morphological features of these samples such as length, color, shape, presence/absence of root/tip, and root status were observed and compared using low-power stereomicroscopes (Leica MZ 12.5 and Zeiss AXIOZoom.V16). When necessary, hairs were mounted on microscope slides in silicon oil for subsequent high-power microscopic examination. The fine morphological characteristics of these hairs (hair diameter, size, density and distribution of the pigment,

shape of pigment aggregations if present, cortical texture, medulla) were compared to those of the reference samples using a high-power comparison light microscope (Leica FS 4000).

In some PSD patients, the comparative morphological hair examination was hampered or failed for two reasons: (1) an irreversible heavy staining of the hair by toluidine blue as a result of the operation procedure changing the color impression and covering the natural pigmentation, (2) the effects of hair strength testing which caused deformation of the hair shaft, and (3) the exquisitely short length of some sinus nest hair left little biological material to compare to the body hair harvested in full length.

So in most of these hairs, a comparison of relevant morphological features such as color or pigmentation was not fully possible.

Statistical considerations

Axial (trajectory) hair testing results of each single hair harvested and tested were used to estimate probabilities of sinus hair to originate from POE, GS, or IGF, respectively. The Bayesian testing framework based on linear discriminant analysis and multiple testing correction is described in detail in the next section. The statistical software R was used to perform analyses and generate plots.

Statistical analysis

The axial strength of every hair was measured six times in sequence (unless it broke or dislocated before). Before the analyses, all strength measurements were transformed to logarithmic scale. To simplify notation, we always refer to the logarithmic scale when mentioning explicit values of "axial hair strength" throughout this paper. In general, the measured strength of a hair slightly changed from measurement to measurement. In order to aggregate several measurements into a single strength value for each hair, we fitted a linear model with hair and measurement number as explanatory variables and used the predicted strength for the first measurement as a "mean strength" for each hair.

Based on these mean strength values, we estimated the probability of every pilonidal sinus hair to originate from POE, GS, or IGF, respectively. We applied a linear discriminant analysis (LDA) model using axial hair strength as the only explanatory variable for this step. We fitted one LDA model for every patient, using POE, GS, and IGF hair samples to train the model and assigning equal prior probability to the possible hair sources (POE, GS, and IGF). We performed nonparametric bootstrap analysis to quantify to estimate 95% confidence intervals for posterior probabilities. The statistical approach is explained in more detail in the Appendix.

We were mainly interested in distinguishing IGF from non-IGF (that is, POE or GS) hair in the pilonidal sinus hair sample. For each of the 62 sinus hair investigated in this study (total sample size over all patients), we tested the null hypothesis $H_{0,\,i}$, "hair i originates from its patient's IGF," against its alternative $H_{A,\,i}$, "hair i originates from its patient's POE or GS" (i indexes the sinus hair, $i = 1, \ldots, 62$).

We calculated posterior odds $P(H_{0,\,i}|\,x)/P(H_{A,\,i}|\,x)$ for these null hypotheses using the posterior probabilities $P(H_{0,\,i}|\,x)$ and $P(H_{A.\,i}|\,x)$ from the LDA models (x summarizes all data in the formulae). Posterior odds larger than 1 are in favor of the null hypothesis ("hair probably originates from IGF"), and posterior odds smaller than 1 in favor of the alternative ("hair probably originates from POE or GS"). Posterior odds can be written as a product of prior odds and the Bayes factor of the hypotheses:

$$O_i = \frac{P(H_{0,i}|x)}{P(H_{A,i}|x)} = B \cdot \frac{P(H_{0,i}|x)}{P(H_{A,i}|x)}, \text{ with } B = \frac{P(x|H_{0,i})}{P(x|H_{A,i})} \text{ the Bayes factor or likelihood ratio.}$$

The Bayes factor does not depend on the choice of priors for null and alternative hypothesis and indicates to what extent the posterior odds are influences by data alone.

When calculating posterior odds for 62 hair samples individually, it is likely to attribute a high probability of originating from POE or GS to at least some hair samples just by chance. We control for multiple testing by correcting priors for $H_{0,\,i}$ using the following procedure proposed by different authors [25, 26, 27]: we consider the joint null hypothesis $H_0 = H_{0,\,1} \cap \ldots \cap H_{0,\,62}$ that *all* sinus hair samples originate from the IGF and assign it a prior probability of 0.5. This is equivalent to assigning each individual null hypothesis $H_{0,\,i}$ a prior probability of 0.9888, or using a prior odds of 88.95. Note that this correction makes a test decision in favor of any alternative harder. On the other hand, rejecting any $H_{0,\,i}$ and therefore also the joint null hypothesis H_0 is a statement in favor of the alternative that at least one sinus hair originates from POE or GS.

Results

Axial hair strength was tested in each of the 20 patients from three body regions' hair and from the sinus nest. In 16/20 of all cavities, the strongest hair present resembles the strength of the POE hair. Please note that the sum of all [n] from POE, GS, and IGF is above 20 due to allowed double characterization (as in line 2, where POE and GS hair are both a possible source of the strongest sinus cavity hair, or in line 5, where all three regions may be the source of the strongest sinus hair) (Table 1).

Table 1 Clinical prima vista hair strength analysis for 20 pilonidal sinus patients, comparing sinus hair strength with other body hair strength (tested regions were POE = protuberantia occipitalis externa; GS = glabella sacralis; and IGF = intergluteal fold)

Hair strength	Total [*n*]	POE [*n*]	GS [*n*]	IGF [*n*]
POE ≥ PSD	5	5	–	–
POE + GS ≥ PSD	1	1	1	–
POE + IGF ≥ PSD	4	4	–	4
IGF ≥ PSD	2	–	–	2
POE + GS + IGF ≥ PSD	6	6	6	6
PSD ≥ POE + GS + IGF	2	–		
Total	20	16	7	12

Thus, minimum one half (16/20) of sinus hair may be originated from the POE, estimated by this method of max. axial strength comparison.

Forensic (comparative) microscopic examination of 20 sinus nest and body region sets revealed no sinus hair macroscopically in 5 patients. In 7 patients, hair was found but the source of hair could not be determined despite best efforts due to short hair fragments or methylene blue staining blurring hair surface characteristics (Table 2; n.a.p.). In the remaining 10 patients, 5 presented with head (POE) hair in the sinus, 5 with hair from the glabella sacralis region, (GS), and 4 with hair from the intergluteal fold (IGF). Thus, in 5 out of 10 sinus nests where morphological hair allocation was possible, hair typical for the POE region could be found.

Figure 1 shows Bayes factors and posterior odds calculated as described in the section "Statistical analysis" for the 62 single sinus hair investigated by axial strength testing in this study. Seven hair (11%) have a Bayes factor < 0.01, and 3 (4.8%) a posterior odds < 0.0001; this means that for 3 hair samples, the probability of originating from POE or GS is 10,000 times higher than the probability of originating from IGF.

Due to the small sample size per patient, our fitted LDA models showed a large uncertainty in posteriors and hence Bayes factors. To account for that, we repeated the calculation

Table 2 Morphological allocation of sinus hair to three body regions along the dorsal sweat crest. Green area: examinations by forensic biologist 1 (CB 1); brown area: examinations by forensic biologist 2 (CB 2). *POE* protuberantia occipitalis externa, *GS* glabella sacralis, *IGF* intergluteal fold, *n.a.p.* no allocation possible

Forensik	POE	GS	IGF	n.a.p.	empty nest
1				1	
2	1				
3	1	1	1		
4			1		
5				1	
6		1			
7				1	1
8			1		
9	1				
10				1	
11		1			
12				1	1
13				1	
14	1				
15				1	
16		1	1		
17	1	1			
18					1
19					1
20					1
CB 1	3	2	3	4	1
CB 2	2	3	1	3	4
Total	5	5	4	7	5

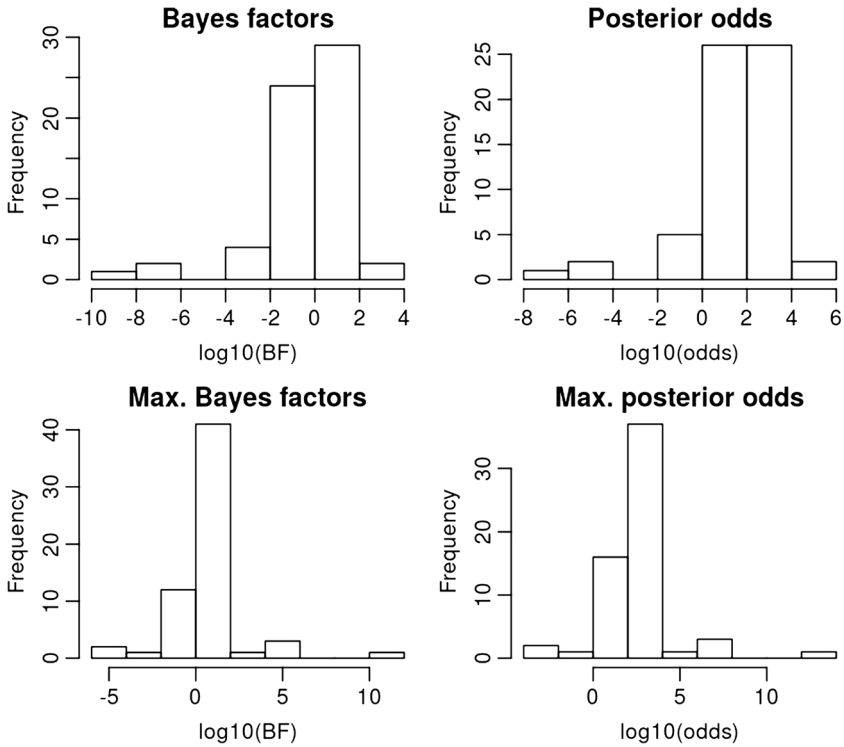

Fig. 1 Upper row: distribution of estimated Bayes factors and posterior odds for the null hypotheses that hair stems from a patient's IGF (one value per pilonidal sinus hair). Lower row: upper limit of the 95% confidence interval for Bayes factors and posterior odds, calculated by bootstrap

of posterior odds for the upper limit of the 95% confidence intervals for the Bayes factors calculated by bootstrap. The resulting distribution of these maximal Bayes factors and posterior odds is shown in Fig. 1. Two hair (3.2%) have a maximal Bayes factor < 0.0001 and a maximal posterior odds < 0.01. This result can be interpreted as follows: for 2 hair, we are 95% confident that their probability of originating from POE or GS is at least 100 times higher than their probability to originate from the IGF region.

Detailed heat maps depicting posterior probabilities of all analyzed hair in all patients can be found in the Appendix.

Putting the heat map results together in one table, the likelihood of POE hair being in the sinus is 10/14, whereas hair from the glabella is likely in 8/14 patients, and intergluteal hair by 9/14. In minimum, one third of the patient hair axially strength tested and analyzed by this bootstrap method, and head hair is the most likely to be in the sinus nest (Table 3).

Discussion

Determining the source of pilonidal sinus nest hair, we used a trifold approach of axial hair strength testing, Bayesian calculation, and criminal biology morphologic comparison.

Hair strength measurements show that 16 of 20 patients had strongest occipital hair which could penetrate the skin most forcefully. Comparative morphological investigations using microscopic aspects give evidence that in 5 out of 14 nests examined with determinable hair, this is highly likely to be hair from the occipital regions—either alone or in combination with hair from other regions. Classification based on LDA models of hair strength measurements shows that with 95% confidence, 2 nests contain hair samples whose probability to stem from POE or GS is at least 100 times higher than to stem from IGF.

Thus, our null hypothesis that "hair in the sinus cavity is originated from the intergluteal fold" can be rejected. Intergluteal hair is not the single source of hair in the pilonidal sinus nest cavity.

Although electron scanning microscopy and new techniques such as "coating" with carbon or gold evolved significantly since the publications by Dahl [8], hair source determination has been proven to be notoriously difficult. Even the FBI acknowledges that "..some people can share the same

Table 3 Heat map results with most likely source of pilonidal sinus hair in $n = 20$ patients. Please note that in $n = 6$ patients, the amount of hair in the sinus was zero or too scarce to calculate reliable probabilities

Most likely hair source	Total	POE	GS	IGF
POE	2	2	–	–
POE and UR	2	2	2	–
POE and IGF	1	1	–	1
GS	1	–	1	–
IGF	3	–	–	3
POE, GS, and IGF	5	5	5	5
Total	14	10	8	9

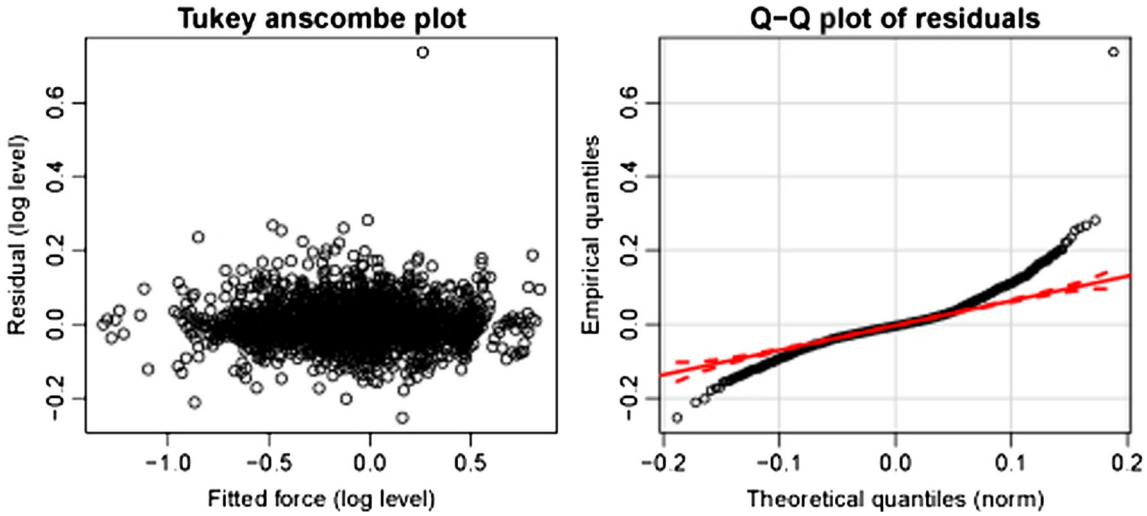

Fig. 2 Tukey-Anscombe plot (residuals vs. fitted values) and residual quantiles vs. normal quantiles. Plots for validation of the linear model used for aggregating repeated hair strength measurements

microscopic characteristics [of hair]," and especially shorter hair segments are impossible to describe [22]. DNA analysis, now used in forensic medicine, is able to distinguish hair from different persons, if the hair root is present. Up to date, it is not able to allocate hair to a certain unknown body region from the identical person. Hair analysis was widely known and praised as a science, but due to several setbacks, its use has been severely downgraded from the 1990s on. This explains why gathering forensic hair evidence is particularly difficult and time consuming. Nevertheless, our forensic partners could gather evidence that in one third of all nests, hair from the occiput was present.

In excess of our results, one third of hairs could not be classified by our forensic biologists. It is redundant to speculate where the short hair fragments that have not been classified due to their short length were coming from—they must have been generated where hair was shaven or cut, producing shorter pieces of hair. This is the fourth reason pointing to the head as a source of pilonidal hair.

Using the maximum strength comparison as well as the statistical analysis, it could again be proven that POE hair can be found within the sinus nest.

It has been long speculated, but never proven, that head hair may play a more significant role in PSD as thought before. Nobody could explain why women without any intergluteal, glabella, or shoulder hair were blighted with pilonidal sinus disease. It has been ignored that young soldiers (even not driving jeeps [7]) do present with PSD, but young lorry driver and bus driver do not. Why is military service linked to a higher PSD risk at young age [1, 24], whereas civil service is not? Short hair needs to be cut more often, which is exactly what is typical for the military environment, leading to thousands of short hair fragments during every dry haircut, with part of it sliding down along the dorsal sweat crest every

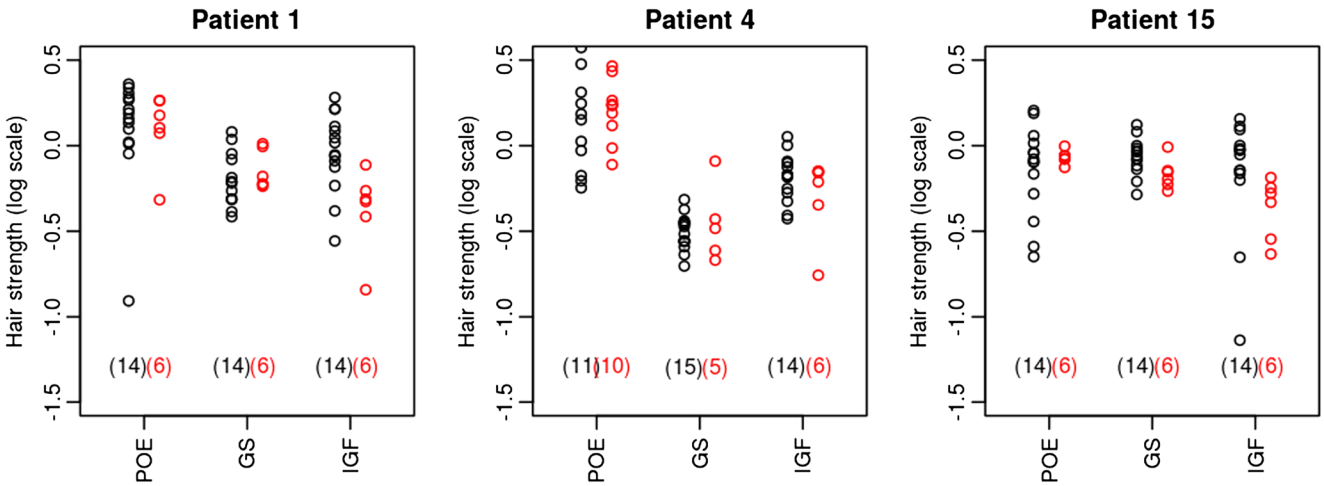

Fig. 3 Comparison of hair strength measurements for randomly harvested hair (black) and hair selected by visible strength (red)

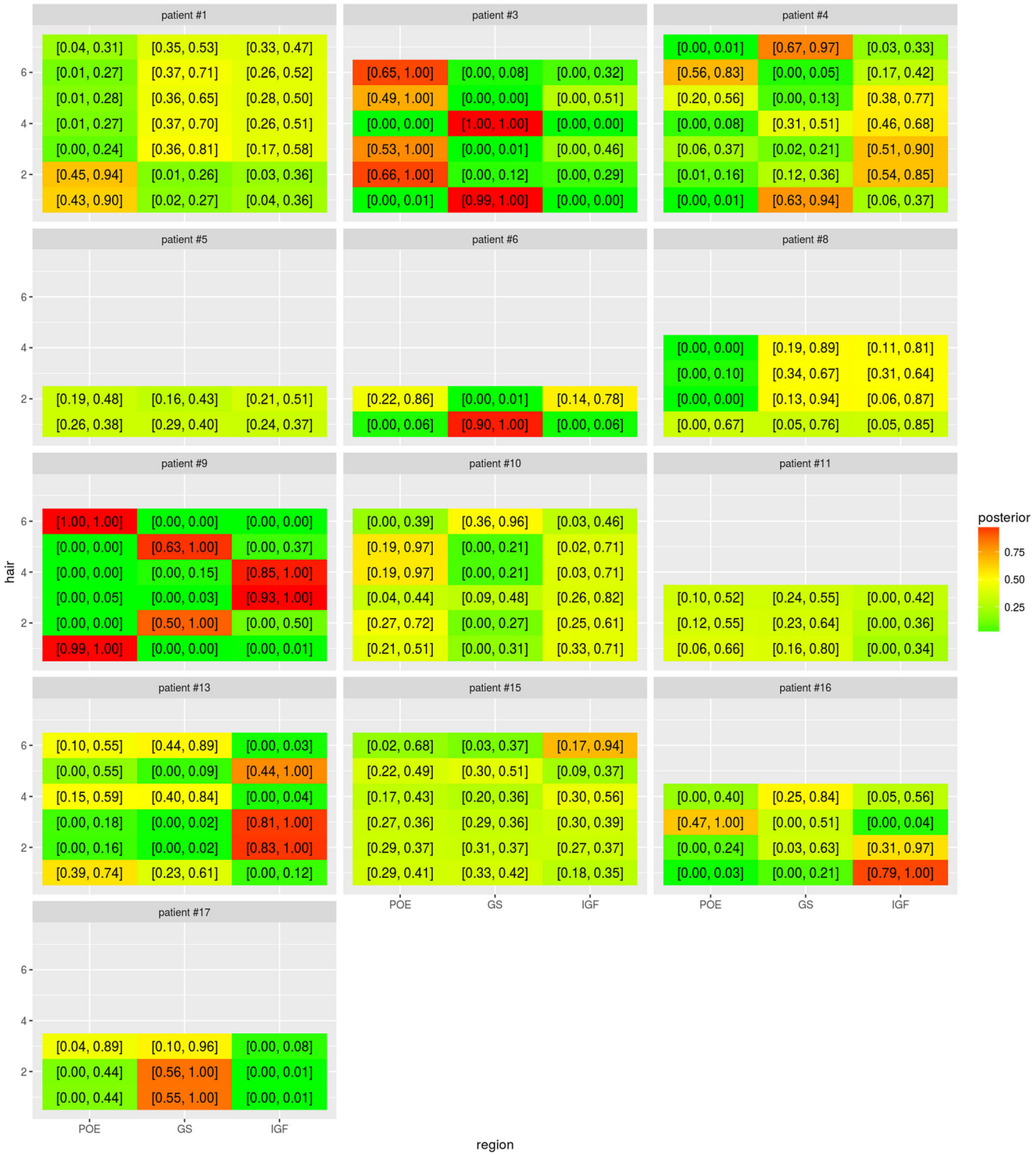

Fig. 4 Posterior probabilities of all classified hair to originate from POE, GS, or IGF, respectively. Numbers in brackets denote the 95% confidence intervals calculated by bootstrap. Red cells refer to a high probability and green cells to a low probability. Patients not shown here did not have a sinus hair sample (patients #7, #12), or no IGF hair sample (patients #2, #14; the null hypothesis that sinus hair stems from IGF cannot be tested in this case). Reading example for patient #1: hair no. 1 is assigned a probability between 43 and 90% of originating from POE, 2 to 27% of originating from GS, and 4 to 36% of originating from IGF

time. Machine cut produces more hair fragments as classical cut with scissors though. Of course, bathing or showering has a positive effect [5, 30], at it removes these short hair fragments before they may be erected and injected into the intergluteal midline. The hygiene argument, often copy echoed but never proven, though is obsolete [11, 15]. Women

have less exposure to short hair fragments, as longer hair is cut less often, leading to longer hair pieces. Due to their weight and being wet, they follow the force of gravity.

The proponents of the folliculitis theory suggest that sebaceous plugs lead to folliculitis [29]. Through the infectious destruction of the skin, hair should be injected into the subcutaneous area and consecutively form a sinus nest. Senapati goes as far and postulates that all sinus on the body may be puncture sinus—except in the intergluteal fold, where it is keratin plug and folliculitis related [25]. It has to be acknowledged that there are exceptionally few sinus patients that are able to report about infectious episodes of the internatal cleft. According to the plug/folliculitis theory, a full-length hair with a root would be the main finding within the sinus nest, which is more the exemption than the rule as we could show [6]. Multiple short hair fragments shorter 2 cm are more than 90% of the sinus hair "population" [10]. And hair pierces into the skin preferentially without root, and with the root decapitated end first, as a recent scanning electron microscopy study of Gosselink and coworkers have shown [16]. Where are these short sharp rootless hair fragments coming from? The only common regularly shaven region *above* the intergluteal fold is the head.

There are some limitations to this study, which are due to the small sample size of 20 hair sets from patients, and not all could be fully identified by the forensic biologists due to their shortness, giving minimal surface characteristics to study. The cohort consists of a northern German population, so data may not be applicable to Mediterranean or other populations outside northern Europe. The biological determination of hair allocation is a difficult and arduous task, which needs a lot of expertise. It is very personal though, but both forensic biologists—examining half of the samples each—came to the same amount of head hair found in the sinus nests.

Today, in a synopsis of three different approaches, all three methods could prove independently from each other that occipital hair is present in the sinus cavity. The amount of occipital hair needs to be clarified. The most pressing question is, if this head hair—as it may exert exquisite stiffness and has razor sharp edges [6]—is the first intruder into the healthy skin, or are the fragments just following a beaten path? Can we reduce the pilonidal incidence (and recurrence rate) by reducing the amount of head hair fragments within the internatal cleft? And what concomitant mechanisms within the skin render youngsters between age 15 and age 25 more vulnerable to pilonidal sinus than elder adults?

In conclusion, we rejected our null hypothesis H_0 that "hair in the sinus cavity is from the intergluteal region" by each of three different approaches. Occipital hair can be found in the nests of pilonidal sinus patients. Its role for the development of the pilonidal sinus now needs to be redefined.

If proven to be causal, this would open excitingly new perspectives for the understanding and prevention of a common and blighting disease of the young.

Author contributions Statistical analysis and calculations: AH, DD
Manuscript editing and interpretation of data: DD, FB, MML, FR, IS, AH, PM
Manuscript writing: DD, AH, DD, JG, FR, FB, MML
Graphic design: DD, AH, FB
Data acquisition: FB, AH, DD, JG, FR

Compliance with ethical standards

The ethics committee of the medical association of Niedersachsen, Berliner Allee 20, 30175 Hannover, Germany (Prof. Dr. med. Andreas Creutzig, chair), fully and unanimously approved the study based on § 15 of the Niedersachsen Medical Association's professional code of conduct.

Conflict of interest The authors declare that there is no conflict of interest.

Appendix. Statistical remarks

Aggregating hair strength measurements

The strength of every hair was measured up to six times in sequence; we used a linear model with the hair as a factor variable and the measurement number as a numerical variable to aggregate multiple measurements into a single strength value for each hair (see the section "Statistical analysis"). The most important assumption of this model, the homoscedasticity assumption, was fulfilled by the data (see Fig. 2, left), but the distribution of the residuals was rather long-tailed and slightly skewed than normal (see Fig. 2, right). Despite this fact, we considered the overall model fit good enough to be used for aggregating multiple strength measurements for hair samples.

Classification of pilonidal sinus hair to regions of origins

We performed linear discriminant analysis (LDA) to statistically attribute pilonidal sinus hair of each patient to their most probable region of origin (protuberantia occipitalis externa [POE], glabella sacralis [GS], or intergluteal fold [IGF], respectively). This proceeded as follows:

1. We fitted one LDA model per patient, using hair strengths from POE, GS, and IGF samples to train the model. The LDA model assumes hair strengths of each region of origin are normally distributed with equal variance, but different means.
2. We classified each pilonidal sinus hair of a patient by calculating its posterior probabilities to originate from POE, GS, and IGF, respectively. This calculation was done based on the normal distributions of hair strength fitted in step 1, and based on the prior probabilities, we assigned to each region of origin; we assigned equal prior

probability to each region of origin (POE, GS, and IGF; "objective" or "uninformed" prior.

3. In order to quantify the uncertainty in the calculated posteriors, we performed non-parametric bootstrap. That means, we resampled hair samples for each patient 999 times (random sampling with replacement), and repeated steps 1 and 2 above for every simulated sample. We used the empirical distribution of the 999 bootstrap posteriors calculated like this to derive 95% confidence intervals for the estimated posteriors.

This analysis has several potentially delicate points:

- Assumption of a model (normally distributed hair strength with equal variance in each region of origin) that may or may not be appropriate. By visual inspection of the distribution of hair strengths in different regions (POE, GS, and IGF), we convinced ourselves that the assumption of this model seems safe.
- Selection bias. In order to reliably fit an LDA model, we would need (random) hair samples from the three regions (POE, GS, and IGF). However, it is not possible to ensure random sampling in practice: hair are chosen and harvested by human researchers. The sampling strategy that was applied for most patients actually contradicted this requirement: in the beginning of the study, hair were sampled based on visual strength, having forensic examination and visual inspection in mind. To estimate the validity of models trained on these non-random hair samples, new samples were collected from three patients (#1, #4, #15), this time making sure selection is as random as possible for a human researcher. We compared these randomly selected hair samples to the older samples selected based on visible strength.

Interestingly, the hair samples collected with the aim of selecting strong hairs from the different regions were on average not stronger than the random samples (Fig. 3). When comparing hair strength from random and non-random sampling with *T* tests, we got unadjusted *p* values of 0.018 (patient #1, IGF; mean hair force from randomly sampled hair is larger), 0.045 (patient #15, GS; mean hair force from randomly sampled hair is larger), and above 0.05 in other cases. After adjusting the *p* values for multiple testing with the Holm correction, the minimal (adjusted) *p* value was 0.160. We concluded that the problem of non-random selection of hair samples can be neglected in practice. It seems that it is difficult to determine the hair strength by eye, and that the samples of supposedly strong hair are indeed representative for the whole range of hair strengths of the patients, which is good for fitting the LDA models.

- Small sample size. Most models are based on only 6 hair from each region (POE, GS, and IGF); the calculated posteriors are therefore not very precise. We performed the non-parametric bootstrap (see above) in order to quantify the uncertainty. As expected, confidence intervals are large for most hair samples, and most pilonidal sinus hair cannot be statistically attributed to a single region of origin with high probability (see also detailed results in Fig. 3).

Classification results for all pilonidal sinus hair samples are depicted in Fig. 4.

References

1. Akinci OF, Bozer M, Uzunköy A, Düzgün SA et al (1999) Incidence and aetiological factors in pilonidal sinus among Turkish soldiers. Eur J Surg 165:339–342
2. Allen-Mersh TG (1990) Pilonidal sinus: finding the right track for treatment. BrJSurg 77:123–132
3. Anderson AW (1847) Hair extracted from an ulcer. Boston Med Surg J 36:74
4. Ardelt M, Dennler U, Fahrner R, Hallof G, et al. (2017) Adolescence is a major factor of sinus pilonidal disease - gender specific based investigation of case development in Germany from 2007 until 2015. Chirurg 88(11):961–967
5. Bolandparvaz S, Moghadam Dizaj P, Salahi R, Paydar S et al (2012) Evaluation of the risk factors of pilonidal sinus: a single center experience. Turk J Gastroenterol 23:535–537
6. Bosche F, Luedi MM, van der Zypen D, Moersdorf P et al (2017) The hair in the sinus: sharp-ended rootless head hair fragments can be found in large amounts in pilonidal sinus nests. World J Surg 42(2):567–573
7. Buie LA (1944) Jeep disease (pilonidal disease of mechanized warfare). South Med J 37:103–109
8. Dahl HD, Henrich MH (1992) Light and scanning electron microscopy study of the pathogenesis of pilonidal sinus and anal fistula. Langenbecks Arch Chir 377:118–124
9. Davage ON (1954) The origin of sacrococcygeal pilonidal sinuses based on an analysis of four hundred sixty-three cases. Am J Pathol 30:1191–1205
10. Doll D, Bosche FD, Stauffer VK, Sinicina I, Hoffmann S, van der Zypen D, Luedi MM (2017) Strength of occipital hair as an explanation for pilonidal sinus disease caused by intruding hair. Dis Colon Rectum 60:979–986
11. Doll D, Luedi MM, Wieferich K, van der Zypen D et al (2015) Stop insulting the patient: neither incidence nor recurrence in pilonidal sinus disease is linked to personal hygiene. Pilonidal Sinus J 1:11–19
12. Doll D, Stauffer VK, Luedi MM (2016) Intra-anal pilonidal sinus disease: a unique diagnosis possibly pointing to the occiput. ANZ J Surg 86:622–625
13. Elliot D, Quyyumi S (1981) A "pennanidal" sinus. J R Soc Med 74: 847–848
14. Evers T, Doll D, Matevossian E, Noe S, Neumann K, Li HL, Hüser N, Lüdde R, Hoffmann S, Krapohl BD (2011) Trends in incidence and long-term recurrence rate of pilonidal sinus disease and analysis of associated influencing factors. Zhonghua Wai Ke Za Zhi 49:799–803
15. Favre R, Delacroix P (1964) Apropos of 1,110 cases of pilonidal disease of coccy-perineal localization. Mem Acad Chir (Paris) 90: 669–676

16. Gosselink MP, Jenkins L, Toh JWT (2017) Scanning electron microscope imaging of pilonidal disease. Tech Coloproctol 21(11): 905–906
17. Karahan Ö, Eryilmaz MA, Torun V, Sevinç B et al (2010) Is the increase in the number of pilonidal sinus surgery normal? Turkish J Surg 26:207–211
18. Klass AA (1956) The so-called pilo-nidal sinus. Can Med Assoc J 75:737–742
19. Lord PH (1970) Unusual case of pilonidal sinus. Proc R Soc Med 63:967–968
20. Mohanna PN, Al-Sam SZ, Flemming AF (2001) Subungual pilonidal sinus of the hand in a dog groomer. Br J Plast Surg 54:176–178
21. Mount LA (1949) Congenital dermal sinuses as a cause of meningitis, intraspinal abscess and intracranial abscess. J Am Med Assoc 139:1263–1268
22. Oien CT (2009) Forensic hair comparison: background information for interpretation. FBI Forensic Case Rev 11
23. Patey D (1971) Pilonidal sinus: a postscript. Lancet 1:245
24. Sekmen Ü, Kara VM, Altintoprak F, Şenol Z (2010) Pilonidal sinus in the army: its incidence and risk factors. Turkish J Surg 26:95–98
25. Senapati A (2012) The management of pilonidal sinus disease. Contemporary Coloproctology Springer-Verlag London Ltd: 67–75
26. Stone HB (1924) Pilonidal sinus: coccygeal fistula. Ann Surg 79: 410–414
27. Tait GB, Wilks JM, Sames CP (1948) Pilonidal sinus in a barber's hand. Lancet 2:121
28. Vaiude P, Dhital M, Hancock K (2011) A true pilonidal sinus in the hand of a sheep shearer. J Surg Case Rep 2011:6
29. Warren JM (1854) Abscess, containing hair, on the nates. Am J Sci 28:113
30. Yildiz T, Elmas B, Yucak A, Turgut HT, Ilce Z (2017) Risk factors for pilonidal sinus disease in teenagers. Indian J Pediatr 84:134–138
31. Zimmerman K (1970) Pilonidal disease. Dis Colon Rectum 13: 330–332