Endogenous Steroid Profiling in the Athlete Biological Passport

Pierre-Edouard Sottas, PhD*, Martial Saugy, PhD, PD, Christophe Saudan, PhD

KEYWORDS

• Androgen • Doping • Steroid • Profiling
• Athlete biological passport

Anabolic–androgenic steroids (AAS) represent a class of steroidal hormones affiliated with the hormone testosterone. Testosterone is produced naturally in the human body and conjugated mainly with glucuronide and sulfate before excretion in urine (phase 2 metabolism). The androgenic effects of testosterone and its prohormones generally are associated with masculanization and virilization, while its anabolic effects are associated with protein building in the body.[1] In power sports, exogenous AAS primarily are used as myotrophic agents to promote muscle mass and strength. Although their efficacy in terms of improved physical function has been debated during decades,[1,2] a comprehensive study by Bhasin and colleagues demonstrated in 1996 that testosterone can act as a performance-enhancing substance when supra-physiological doses are administered.[3] Exogenous AAS also are known to be used in endurance sports for improved recovery. Endurance athletes favour low (to limit myotrophy) but frequent doses for replacement levels. Indeed, overtraining-induced stress can upset the balance between anabolic and catabolic states of the hormones of the endocrine system.[4] Some endurance athletes may find in synthetic AAS an ergogenic supercompensating agent for sustained testosterone concentrations and, in turn, a performance-enhancing substance to allow more intense training sessions. In addition, it has been shown that testosterone not only plays an important role in muscle metabolism during the regeneration phase after physical exercise, but also seems to increase the ability of the muscle to refill its glycogen storage through an increased activity of the muscle glycogen synthetase.[4,5]

This work was supported by Grant Number R07D0MS from the World Anti Doping Agency.
Swiss Laboratory for Doping Analyses, University Center of Legal Medicine, West Switzerland, Chemin des Croisettes 22, 1066 Epalinges, Switzerland
* Corresponding author.
E-mail address: Pierre-Edouard.Sottas@chuv.ch (P.-E. Sottas).

Endocrinol Metab Clin N Am 39 (2010) 59–73
doi:10.1016/j.ecl.2009.11.003 **endo.theclinics.com**
0889-8529/10/$ – see front matter

Intake of exogenous AAS is not the only way to produce a sustained rise in testosterone levels. Various indirect steroid doping strategies produce the same effect. This includes, among others, estrogen blockade by estrogen receptor antagonists (antiestrogens) or aromatase inhibitors.[6] Although these two classes of estrogen blockers differ in their pharmacologic action, both are known to stimulate sustained increases in endogenous luteinizing hormone secretion, and, successively, increases in blood testosterone concentrations. In particular, estrogen blockade in men is known to produce elevations of testosterone concentrations at a level sufficient to produce ergogenic and performance-enhancing effects.[6]

MARKERS OF STEROID DOPING

As of today, triple quadrupole mass spectrometry cannot distinguish between pharmaceutical and natural testosterone based on the mass spectrum. In the 1980s, pursuant to the work of Donike and colleagues, an authorized upper limit of 6.0 for the testosterone over epistestosterone (T/E) ratio was introduced to deter testosterone administration.[7] Because epistestosterone is only a minor product of the metabolism of testosterone and does not increase after exogenous testosterone administration, the net effect of the latter is an increase in the T/E ratio.[8] Testosterone and epistestosterone levels in urine specimens commonly are measured in antidoping laboratories by gaz chromatography–mass spectrometry (GC/MS) after deconjugating the glucuronide moiety by enzymatic hydrolysis (β-glucuronidase) and derivatization (trimethylsilylation).[9,10] Alternatively, testosterone and epistestosterone can be measured directly using high-performance liquid chromatography (HPLC)/tandem MS.[11]

The T/E has been the first widely used indirect marker of doping with anabolic steroids, with a discrimination principle not based on the distinction between the exogenous substance and its endogenous counterpart, but rather on the effect induced by the intake of the exogenous substance on some selected biological markers. Although the value of evidence provided by population-based limits on biomarkers generally can be considered as being not useful from a forensic perspective,[12] a T/E ratio greater than 6.0 nevertheless was adopted as proof of steroid doping in 1982. Unsurprisingly enough, it was put forth a few years later that some individuals were shown to produce naturally elevated T/E.[13] Since then the T/E ratio mainly has been used as a screening test, with any positive result requiring a subsequent confirmation analysis by GC/C/IRMS. GC/C/IRMS allows measurement of slight differences in $^{13}C/^{12}C$ ratio of testoserone metabolites. Discrimination between pharmaceutical and natural testosterone is possible, because hemisynthetic testosterone is known to display a different ^{13}C content than its human counterpart produced by means of cholesterol metabolism.[14] GC/C/IRMS has become an indispensable tool in antidoping laboratories for the determination of synthetic AAS in urine samples, despite the fact that the method is not sensitive to indirect androgen doping.

LONGITUDINAL STEROID PROFILING

Whereas it already was known in the 1990s that subject-based reference ranges are much reliable than population-based reference ranges for androgens[15] and that individual T/E values do not deviate from the mean value by more than 30%,[8] it has only been recently that a method was proposed to take into account formally these characteristics.[16] Based on empirical Bayesian inferential techniques for longitudinal

profiling that also are used for cancer screening,[17] the test progressively switches the focus from comparison with a population to the determination of individual values. Interestingly, this test is neither a purely population-based nor a purely subject-based approach, but an intermediate approach that makes the best decision in function of the between- and within-subject variance components of the marker and actual individual test results. At each moment in the course of data acquisition, it is possible to predict expected values for the markers and to define individual limits for a desired specificity (assuming a nondoped population). From a mathematical point of view, individualization of the expected values of the marker corresponds to the nullification of its between-subject variance component. Using the athlete as his own reference is particularly interesting when the marker presents a low ratio of within-subject to between-subject variations. In a population composed of male Caucasian athletes, this ratio has been estimated to be as low as 0.04 for the T/E.[18] Such a low ratio already questions the pertinence of a population-based threshold (fixed at 4.0 today)[19] for the T/E ratio.

In addition to general descriptive statistics, the sensitivity and specificity of various methods of interpretation applied to the T/E marker also have been evaluated empirically.[20] In detail, the specificity was estimated from 432 urine samples withdrawn from 28 control subjects, the sensitivity from 88 urine samples collected in a clinical trial at a maximum of 36 hours after administration of a pill of 80 mg of undecanoate testosterone. A population-based limit fixed at 4.0 for the T/E ratio returned 24 false positives (24/432 = 5.6%) for 34 true positives (34/88 = 39%), that is a positive predictive value (PPV) of only 59% on that set of data. A high PPV is particularly important in antidoping, because the PPV indicates the true probability of an athlete doped in case of a positive outcome. A PPV of 59% indicates that when the test returns a positive result, there is 41% chance of a false positive. On the same T/E data, the empiric Bayesian test for longitudinal data returned two false positives (2/432 = 0.5%, for a theoretical specificity of 99%) for 51 true positives (51/88 = 58%), that is a PPV of 96%. Still better results were obtained when some heterogenous and external factors were taken into account. In conclusion, these numbers confirmed the inefficiency of a unique and inflexible threshold for the marker T/E. Actually, from a forensic perspective, it can be shown that the value of the evidence given by the rule "T/E > 4.0" can be considered as being not useful.[12,21]

Every year, the World Anti-Doping Agency (WADA) publishes some statistics on the number of adverse analytical findings (AAFs) and atypical findings (ATFs) reported by antidoping laboratories. AAS represent by far the family of substances that lead to the highest number of AAFs and ATFs. For example, for 2007, the numbers are the following: 223,898 A samples were analyzed, for a total of 4,402 AAFs (1.97%), from which 2,322 (47.9% of all AAFs, 1% of all tests) were for AAS. Accordingly, it often is claimed that AAS, and testosterone in particular, represent the most abundant misused substances in elite sports. However, because all tests returning a T/E value higher than 4.0 are reported as an AAF (as an ATF since 2008) and knowing that a significant proportion of male athletes should present naturally higher values than 4.0,[20] one cannot exclude that most AAFs (or ATFs) obtained with the "T/E > 4.0" rule are false positives. This happens when the prevalence of steroid doping is low and when the test has been applied many times. Given that GC/C/IRMS analysis is particularly costly, the financial waste to apply a "T/E > 4.0" rule can be estimated at about $1 million. This unnecessary financial burden does prejudice expenditure on other more efficient investigations and the credibility of the antidoping movement in general.

STEROID PROFILE

While the terminology steroid profiling is used in the literature to denote a follow-up over time, a steroid profile encompasses concentration levels of endogenous steroids in urine and their respective ratios. Steroid profiles are employed widely in endocrinology to detect enzyme deficiencies or adrenal problems.[22] In antidoping laboratories, the urinary steroid profile usually includes the concentration levels of

Testosterone
Testosterone's inactive epimer, epistestosterone
Four testosterone metabolites, androsterone, etiocholanolone (Etio), 5α-androstane-3α,17β-diol (α-diol), and 5β-androstane-3α,17β-diol (β-diol)
A testosterone precursor, dehydroepiandrosterone (DHEA)

The following cut-off concentration levels of endogenous steroids equivalent to the glucuronide: testosterone >200 ng/mL, epistestosterone >200 ng/mL, androsterone >10'000 ng/mL, Etio >10'000 ng/mL and DHEA >100 ng/mL are considered as putative markers of steroid doping.[19] In contrast to absolute steroid concentrations, ratios such as T/E, androsterone (A)/Etio, A/T, α-diol/E and α-diol/β-diol are robust to circadian rhythm or changes in physiologic conditions such as exercise workload for athletes.[23] On the other hand, these parameters may be altered significantly according to the administered steroid and its application mode.

Although glucuronide conjugates up to now have been the preferred means of evaluating excretion of androgens, there is a high potential of improvement in the development of additional markers of steroid doping through other phase 2 metabolites, such as testosterone sulfate. Methods based on HPLC/MS have been developed for that purpose.[11,24,25] Introduction of sulfoconjugates, with biomarkers such as the ratio testosterone sulfate/ epistestosterone sulfate (Ts/Es), or testosterone glucuronide/testosterone sulfate (Tg/Ts), or (Tg+Ts)/(Eg+Ts) may help develop a more sensitive test for AAS abuse in the future.

HETEROGENEOUS FACTORS

Heterogeneous factors refer to the factors specific to an individual that are known to have an influence on a biomarker. For example, sex and age are well-known heterogenous factors used in the evaluation of a steroid profile.

It long has been known that urinary testosterone glucuronides present a bimodal distribution,[9] this effect being particularly marked between Caucasian and Asian populations.[26–28] It only was recently, however, that it was demonstrated that the significant differences observed in testosterone glucuronide excretion are associated with a deletion mutation in the UDP-glucuronide transferase 2B17 (UGT2B17) gene.[29] This discovery has important implications for doping tests. For example, when subjects deficient in the UGT2B17 gene (del/del) receive exogenous testosterone, it has been shown that their T/E ratio does not rise significantly, remaining well below current threshold at 4.0.[20,30] This suggests that the knowledge of genetic differences in metabolism and excretion is important in the evaluation of urinary steroid profiles.[2]

The understanding of the genetics of androgen disposition has grown quickly. Recent studies have shown, among others, that the cytochrome P-450c17alpha (CYP17) promoter polymorphism may partly explain high natural T/E ratios.[31] Also, the lack of the UGT2B17 enzyme may be compensated for by an increase in UGT2B15 transcription.[31] In addition, it has been shown that epistestosterone does not present a bimodal distribution because UGT2B17 does not glucuronidate E while

UGT2A1 conjugates testosterone and epistestosterone similarly,[25] and, finally, that testosterone is primarily glucuronidated by UGT2B17 while epistestosterone is primarily glucuronidated by UGT2B7.[25] The latter result is particularly interesting to understand the large between-subject variations of the T/E ratio, with low values of T/E (significantly < 1) expressed by subjects deficient in the UGT2B17 gene, and high values of T/E (significantly > 1) expressed by subjects deficient in the UGT2B7 gene. Additional discoveries in the genetics of androgen disposition are expected in the coming years, with genes involved in phase 1 drug metabolism, such as the P450 gene (CYP) family, and in phase 2 metabolism, such as the glucuronosyltransferases (UGT) and sulfotransferases (SULT) gene families.

A recent study by Xue and colleagues has shown that the UGT2B17 gene presents an unusually high degree of geographic variation, with a high frequency of the gene in most African populations, intermediate frequency in Europe/West Asia, and low frequency in East Asia.[32] Interestingly, an impressive worldwide map of the distribution of the UGT2B17 gene has been published for more than 30 ethnic groups. This high variability in the frequency of the UGT2B17 also has been confirmed in sports, in a study with five different ethnic groups of professional soccer players.[33]

All these studies have shown a large heterogeneity in androgen disposition. From a mathematical point of view, a large heterogeneity is expressed through a large between-subject variance of the marker. Introduction of genotyping information of the athlete, or of the frequencies of the genes in function of the ethnic origin of the athlete, can remove the part in the between-subject variance that originates from these differences.[34] These studies confirm, again, that unique and nonspecific thresholds on markers of steroid doping are not fit for indicating AAS misuse.

EXTERNAL CONFOUNDING FACTORS

Several confounding factors must be considered when a longitudinal steroid profile has to be interpreted. For that purpose, it is worth mentioning the review of Mareck and colleagues describing the factors known to exert an influence on the steroid profile.[23] The influence of some pharmaceutical preparations and the potential influence of microorganisms and bacterial activities in urine samples were reviewed. In particular, it is relevant to outline that the consumption of ethanol at dosages higher than 1 g/kg bodyweight may lead to a significant increase of the T/E ratio.[35] Similarly, the elevation of this ratio also can be observed upon application of oral contraceptives owing to suppression of epistestosterone excretion.[36] In contrast, the application of ketoconazole is known to lead to an inhibition of stereoidogenesis and subsequently to result in a suppressed urinary profile and significant variation of the T/E ratio.[37]

STANDARDIZED PROTOCOLS FOR STEROID PROFILE DATA

In a medico–legal setting, it is the burden of the testing officials to demonstrate the validity of the presented evidence. In that context, the measurement of a steroid profile must follow standardized procedures based on justifiable protocols. Such compliance is necessary to control analytical uncertainties. This is particularly important in steroid profiling, because it is essential to quantify the expected variations of the markers. For GC/MS in particular, the effects of some technical parameters such as inhibition of hydrolysis, incomplete derivatization, or matrix issues must be under control.[23] For example, if two laboratories do not have the same limit of quantification (LOQ) for the concentration of a steroid, the analytical uncertainty will be different for the measurement of a concentration close to the LOQ, with, in turn, a within-subject variance of the markers that may change from one sample to the next. Therefore,

a standardization of the protocols with all laboratories hanging to an external quality control system is an essential condition for a forensic evaluation of steroid data.

BAYESIAN INFERENCE FOR THE EVALUATION OF INDIRECT EVIDENCE

The causal relationship between doping (the cause) and the induced modification in the steroid profile (the effect) can be formalized and graphically represented by a causal network (**Fig. 1**A). The goal is to establish whether an athlete is doped by examining his steroid profile. This type of problem goes against the causal direction, and the only logic that may apply here is Bayesian reasoning.[38] For example, if an athlete takes a oral dose of synthetic testosterone (the cause), the value of his T/E ratio

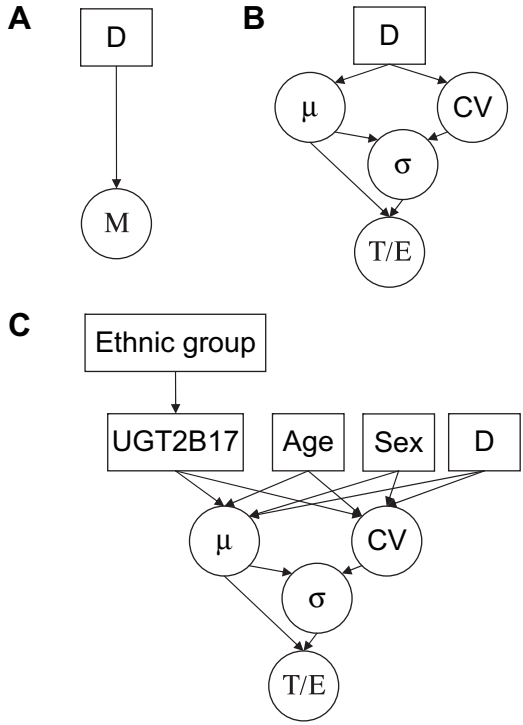

Fig. 1. Bayesian networks for the evaluation of the evidence with markers of steroid doping. Each rectangle presents a discrete variable, each circle a continuous variable, each arrow a causal relationship. (*A*) D represents the doping state of the athlete, M the marker. Doping (the cause) has an effect on the marker (the effect), and the goal is to know in which doping state the athlete is in light of the result of the marker M. This problematic goes against the causal direction, and only Bayes' theorem handles this point. (*B*) A longitudinal approach for the T/E can be modeled by making explicit the expected mean and coefficient of variation of the sequence of T/E values. A log-normal distribution is assumed for these two variables. (*C*) Addition of the heterogeneous factors age, sex, and UGT2B17 genotype. If the athlete's genotype is unknown, the prevalence of the UGT2B17 gene can be used instead in function of the ethnic origin of the athlete. Similar causal networks are used in genetics to represent so-called genotype-phenotype maps. (*Data from* Xue Y, Sun D, Daly A, et al. Adaptive evolution of UGT2B17 copy-number variation. Am J Hum Genet 2008;83:337–46. Rockman MV. Reverse engineering the genotype–phenotype map with natural genetic variation. Nature 2008;456:738–44.)

will increase (the effect). If a model that links cause and effect is available, Bayes' theorem can be used to follow the direction that is opposite to that of causality and to determine whether an increase in the T/E may be the result of doping or is caused by natural variations. In such a Bayesian network (see **Fig. 1**A), D is a discrete variable describing the doping state of the athlete, and M a continuous variable representing the result of a measurement of a marker of steroid doping (such as the T/E). According to Bayes' theorem, the causal relationship between M and D can be written as follows:

$$P(D|M) = \frac{P(M|D) \cdot P(D)}{P(M)} \quad \textbf{(1)}$$

where the formulations are given as probabilities. $P(D|M)$ represents the probability of being in state D as a function of the value of marker M, $P(M|D)$ the probability to measure the result M knowing that the athlete is in state D. For example, if M is T/E and $D = 0$ represents the nondoped state, $P(\mathrm{T/E}|D = 0)$ is well described by a log-normal distribution with geometric mean of 1.40 and geometric standard deviation of 1.78 for a population of male athletes not deficient in the UGT2B17 gene. The advantage of a Bayesian approach resides in the possibility to use the conditional probability function $P(M|D)$ to determine $P(D|M)$, because a cause-to-effect relationship is much easier to establish than the reverse effect–cause relationship. The cause-to-effect relationship is typically built from data obtained in clinical trials, with control subjects to obtain samples of class $D = 0$, and volunteers to which a doping product has been administered to obtain samples of class $D \neq 0$. Dealing with reference populations of doped athletes requires, however, some assumptions on the type of doping (eg, type of substance or method, doses, toxicocinetics of the substance). Formally, this can be taken into account by means of multiple states of doping $D \in \{1, 2, 3\ldots\}$, with the state $D = i$ specific to the type of substance and type of treatment used to obtain the data on which $P(M|D = i)$ has been developed. Assuming for the sake of simplicity and without loss of generality a unique state of doping $D = 1$, the conditional probability $P(M|D = 1)$ plays an important role in the odds form of Bayes' theorem:

$$\frac{P(D = 1|M)}{P(D = 0|M)} = \frac{P(M|D = 1)}{P(M|D = 0)} \cdot \frac{P(D = 1)}{P(D = 0)} \quad \textbf{(2)}$$

in which the first ratio in the right side of the equation is the so-called likelihood ratio, also known as the weight of evidence in forensics.[21] A likelihood ratio with a value greater than 1 leads to an increase in the odds (to favor the state $D = 1$), while a likelihood ratio with a value less than 1 leads to decrease in the odds (to favor the state $D = 0$).

In antidoping, it is not uncommon to have a decision rule based on a high threshold of specificity of a test or marker, with the underlying assumption that the athlete is part of a population composed of nondoped athletes only. This amounts to defining $\Pr(D = 0) = 1; \Pr(D \neq 0) = 0$ and this removes the true essence of a Bayesian approach. In that situation, only models for nondoped control athletes are required, not for doped athletes. To use a decision rule based on the specificity of the marker has large implications on the logic to evalute the value of evidence. In particular, if proper precautions are not taken, this logic can lead to the so-called false-positive fallacy,[21] a special case of the more general prosecutor's fallacy that results from misunderstanding the notion of multiplicity of tests. An increase in the number of tests on a population composed of nondoped athletes causes an increase in the probability of obtaining a false positive. For example, the limit at 6.0 introduced in 1982 for the T/E ratio was founded on the fact that nobody from a large number (say N) of control

subjects had presented a so high T/E value at that time. It was thus believed that the specificity of this rule was close to 100%. It can be shown by Bayesian statistics, however, that if a number N of control subjects did not present a T/E greater than 6.0, the mean of the distribution of the expected specificity of the classification rule "T/E > 6.0" is $2N/(2N+1)$. Consequently, one can expect one false positive in average if the test is applied twice (ie, $2N$) as many times the number (ie, N) it was applied during its validation. Even though these are only theoretical considerations and that it remained possible at that time that no athlete would have ever presented a T/E greater than 6.0 (only empirical evidence is discussed here), care should have been taken because of the multiplicity of antidoping tests. Mathematically, the false-positive fallacy generally results from the fallacious transposition of the conditional distributions:

$$P(D = 1|M) = 1 - P(M|D = 0)$$

To base the decision solely on a threshold of specificity is not necessary if an estimate of the prevalence of doping is available. Although it may be thought at first sight that only a test able to easily identify drug cheats can be used to estimate the prevalence of doping, it has been shown recently that it can be accurately determined if several conditions are met.[12,39] The idea is to compare the test results of a population of athletes, such as when all athletes participating to the same competition are tested just before that competition, with reference populations of nondoped and doped athletes. For example, if the number of athletes is sufficiently large, the maximal difference between the empirical cumulative distribution function (ECDF) constructed from the values of the marker M on the tested population and the CDF of a reference population of nondoped athletes (ie, the CDF computed from $P(M|D = 0)$), represents a minimal estimate of the prevalence of doping in the tested population. This estimate then can be used as prior distribution $P(D)$ in Equation 1 (or via prior odds in Equation 2) for all athletes who participated to that competition. Then, for each athlete individually, it is possible with Bayes' rule to change the prior $P(D)$ in receipt of the individual test result M and the model $P(M|D)$ to obtain the posterior distribution $P(D|M)$ that represents the true probability that the athlete doped. The same logic is possible with the odds form of Bayes' theorem, with the likelihood ratio computed from the individual test result M updating the prior odds estimated from the group of athletes, to obtain the posterior odds on which a decision rule can be implemented.

Independent of the probability distribution on which the decision rule is applied, a Bayesian logic remains essential to take into account other variables (other than doping) in a natural manner. The Bayesian network of **Fig. 1**B is a graphical representation of the model described previously for longitudinal steroid profiling. In more detail, in this empirical two-level hierarchical Bayesian model, the variables μ and CV are unobservable variables with prior distributions assumed to be log-normal. These distributions are updated progressively on receipt of new test results. Bayesian inference permits the move from prior (pretest) to posterior (post-test) distributions, based on the outcome of the test result (this process goes against the causal direction). Then, the posterior distribution of the variables μ and CV can be used to define the expected values of the marker M for a next test (following this time the causal direction). The whole process is repeated, with posterior distributions obtained from the previous test becoming the prior distributions for a new test. When the number of tests is large, the distributions of μ and CV degenerate to the parameters specific to the athlete. Also, further techniques have been proposed to handle the lack of

independence between two successive values,[17,40] a situation that may occur when two urine samples are collected in a too short a period of time.

Similarly, both heterogenous and confounding factors can be modeled in a natural manner in that framework. **Fig. 1**C shows a Bayesian network with heterogenous factors age, sex, UGT2B17 genotype, and ethnic origin for the evaluation of longitudinal T/E data. The strength of that approach relies in its flexibility to integrate sound scientific knowledge. For example, the relationships between the variables ethnic origin and UGT2B17, and between UGT2B17 and μ can be obtained from literature data.[20,25,29,34]

ATHLETE STEROIDAL PASSPORT

The use of indirect markers of doping has a long history, but it is only recently that this testing paradigm has matured into what is called today the Athlete Biological Passport (ABP). An ABP is an individual electronic document that stores any information valuable for the interpretation of indirect evidence of doping. The fundamental principle of the ABP is the monitoring over time of selected biomarkers to reveal the effects of doping. The more elaborated module of the ABP is the Athlete Haematological Passport,[39] which aims to identify blood doping with biologic markers of an altered erythropoiesis.

All ideas and concepts discussed in the precedent paragraphs on the markers of steroid doping can form the basis of an Athlete Steroidal Passport (ASP). Steroid profiling in an ASP appears especially appropriate, because all of these markers are known to present a low ratio of within-subject to between-subject variations.

Fig. 2 shows the results of an ASP for a male adult Caucasian athlete not deficient in the gene UGT2B17, for the marker T/E, T/A, A/Etio and α-diol/β-diol. This athlete has been tested 10 times, with all measurements performed in the authors' laboratory. The center lines represent the actual test results, the upper and lower lines the limits found for a specificity of 99% with the Bayesian network of **Fig. 1**C. For the T/E, initial limits are [0.24 6.88], meaning that only one individual out of a population of 100 male adult athletes not deficient in the gene UGT2B17 should present a value out of this range in average. The knowledge of the UGT2B7 genotype (unknown for this athlete) may still lead to more specific limits, in particular to avoid false positives (with a result higher than 6.88) if the athlete is shown to be deficient in that gene. With the first test result equal to 0.94, the reference range for the second test become 0.34 to 2.44, meaning that only one individual out of a group of 100 male adulte athletes not deficient in the gene UGT2B17 who have shown an initial value of 0.94 must fall out of this range in average. The last range of 0.51 to 1.28 represents the expected range for that athlete if tested an 11th time. With an upper value of 1.28, the athlete has an insignificant margin to monitor his profile with low doses of testosterone.

Fig. 3 shows the T/E sequence of a professional athlete tested 13 times in 1 year. In the upper left insert are represented the limits found with the empirical Bayesian approach without any other knowledge than the athlete is a Caucasian male. The sequence has 11 tests with T/E values significantly inferior to 1.0, and two outliers at 1.2 and 1.5. For example, the value at 1.2 is at the 99.9996 percentile of the distribution of expected values (this is a theoretical consideration, because the approach has not been validated empirically for a specificity higher than 99.9%). The value at 1.2 is abnormal at a level similar to the level that would have been achieved with an initial T/E value at around 22. If the two values at 1.2 and 1.5 are removed from the interpretation (upper right insert of **Fig. 3**), the sequence does not present any suspect

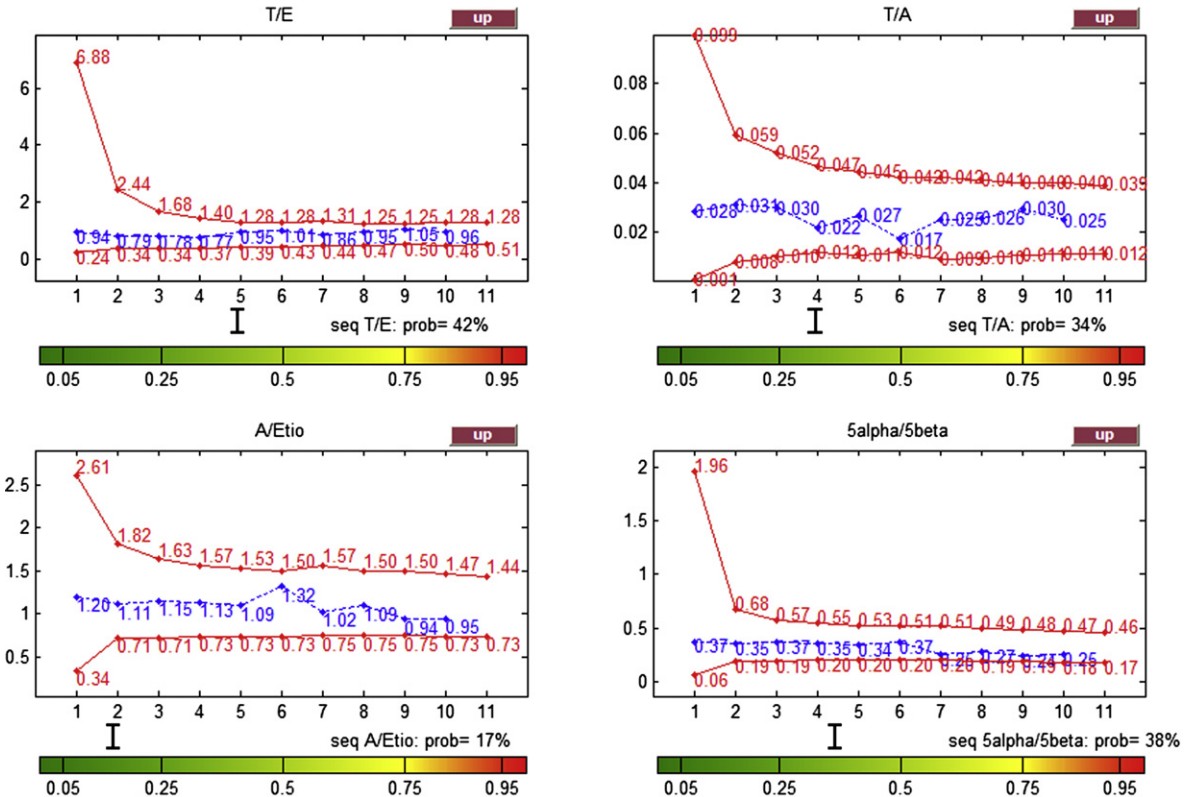

Fig. 2. Steroid profiling for a male Caucasian athlete for the markers T/E, T/A, A/Etio, and α-diol/β-diol. This figure has been obtained from the software Athlete Biological Passport (LAD, Lausanne, Switzerland). This software is available upon request to any antidoping organization.

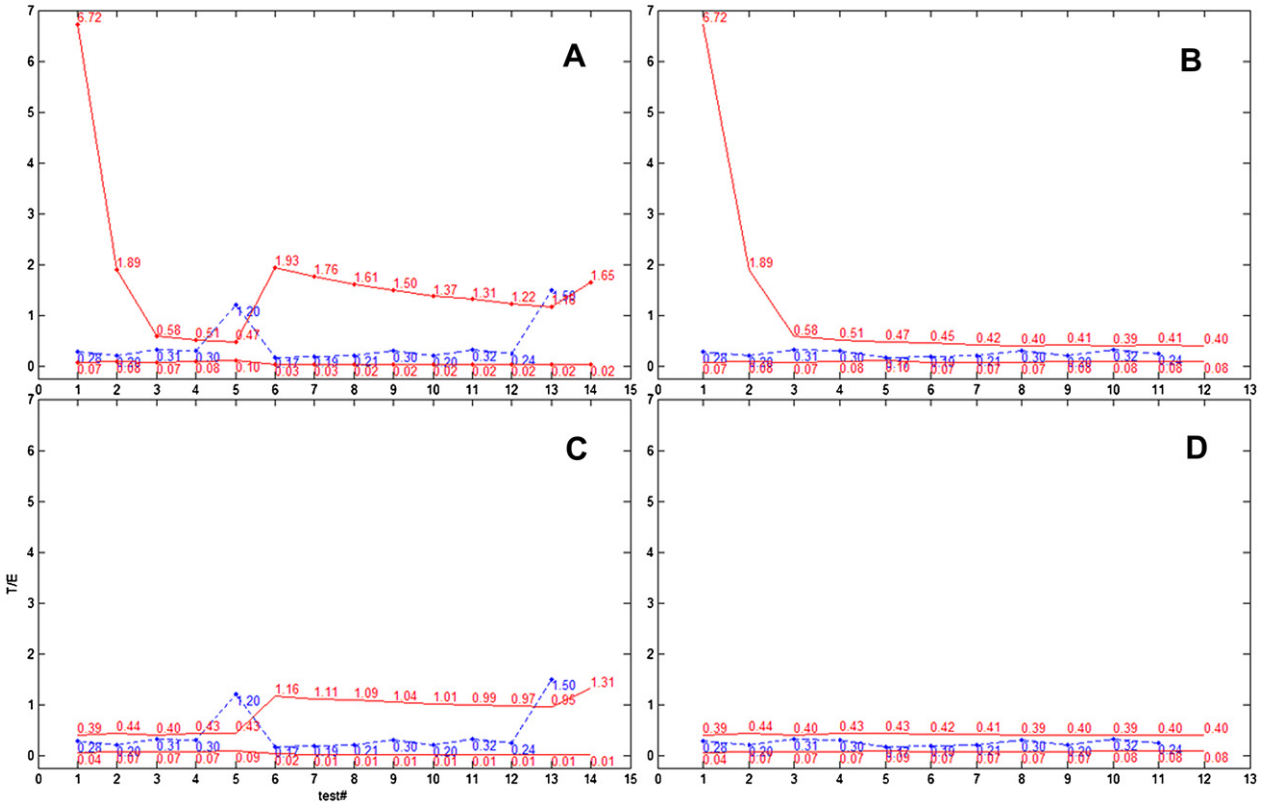

Fig. 3. Longitudinal T/E data for a professional athlete tested 13 times, as well as individual limits for a specificity of 99%. There are 11 tests with values significantly inferior to 1.0, and two outliers at 1.2 (fifth test) and 1.5 (13th test). Even though the athlete's UGT2B17 genotype is unknown, it can be inferred from his phenotype, with a probability of 99% of a full deletion in the UGT2B17 gene (see text). (*A*) All 13 test results are represented with individual limits obtained without any a priori on the UGT2B17 genotype. (*B*) Without the two outliers and without any a priori on the UGT2B17 genotype. *(C)* All 13 tests results and assuming a full deletion in the UGT2B17 gene. (*D*) Without the two outliers and assuming a full deletion in the UGT2B17 gene.

value anymore. With the Bayesian network of **Fig. 2**C, it is possible to infer the posterior probability (or posterior odds) that this athlete is deficient in the UGT2B17 gene in function of his test results, that is to infer the athlete's genotype from his phenotype.[41] With parameters given elsewhere,[16] the likelihood ratio is equal to 0.00092 (odds 1:1080) in favor of a deficiency of the gene. With prior odds of 9:1 in favor of the presence of at least one allele for Caucasians,[20,32] the posterior odds are 1:120 in favor of a deficiency of the UGT2B17 gene for that athlete, that is a probability higher than 99%. The results obtained when it is assumed a priori that this athlete is deficient in the UGT2B17 gene are shown in the lower left (full sequence) and lower right (without the two outliers) for comparison. Interestingly, in that case, the amount of between-subject variance removed by the (assumed) knowledge that this athlete has a full deletion of the UGT2B17 gene is similar to the amount of variance removed by the knowledge of about four basal values.

OUTLOOK OF AN ATHLETE STEROIDAL PASSPORT

The ASP represents the mature product of the development of biological markers of steroid doping. Much progress has been accomplished from the discovery and implementation of the T/E ratio at the beginning of the 1980's to the understanding of the implications of genetic differences in the steroid profile of an athlete. In particular, steroid profiling remains the fundamental principle of the ASP, allowing the removal of the largest part of the variations of the markers. Also, because the compliance to different analytical protocols leads to different expected variations, it is essential that these protocols are an integral part of the ASP.

The benefits of adopting the ASP concept are far-reaching, with multiple applications possible. First, the ASP can be used to target athletes for the GC/C/IRMS test with much better efficiency than it is performed today. Secondly, unusually large disparities found in an ASP may alert officials of doping or a medical condition requiring closer examination. In both cases, the sports authorities have a good reason to withdraw the athlete from competing for a short period, typically 2 weeks. Because fair play and athletes' health protection are fundamental in any antidoping program, the authors strongly believe that the benefits of adopting such a no-start rule would be huge. Also, with a no-start rule, an athlete can use his passport to attest his fair play by means of normal longitudinal profiles of biomarkers. Indeed, whereas a negative outcome of an antidoping test does not necessarily prove that the athlete is clean (because of the low sensitivity of some tests performed at one unique moment in time), the presentation of a passport at the beginning of a competition can ensure that the athlete will participate close to his natural, unaltered physiologic condition. It may be argued that today markers of doping do not have a perfect discrimination aptitude, but the limits set by the passport let a small margin for doping, so that finally the advantages to dope become outnumbered by the risks. Also, contrary to direct tests that must be developed and validated for each new doping substance, a marker is validated and introduced in the ABP once and for all. This means that today's markers only can gain in sensitivity in the future. In particular, it is probable that today's markers are already sensitive to future generations of doping substances (for example to all future substances that will aim to increase testosterone concentrations), whereas the sensitivity of today's direct tests to future substances is far from being guaranteed. Third, when unusually large disparities have been found, the ASP should be reviewed by a panel of experts to determine their origin. When it is much more likely that the detected abnormality originates from doping than from a medical condition or other external cause (eg, confounding factor, multiple testing), the information stored in

an ASP can be sufficient to launch a disciplinary procedure against the athlete. This reviewing process typically can be performed during the short withdrawal of the athlete if a no-start rule has been implemented.

REFERENCES

1. Kicman AT. Pharmacology of anabolic steroids. Br J Pharmacol 2008;154:502–21.
2. Bowers LD. Testosterone doping: dealing with genetic differences in metabolism and excretion. J Clin Endocrinol Metab 2008;93:2469–71.
3. Bhasin S, Storer TW, Berman N, et al. The effects of supraphysiologic doses of testosterone on muscle size and strength in normal men. N Engl J Med 1996; 335:1–7.
4. Urhausen A, Kindermann W. The endocrine system in overtraining. In: Arren MP, Constantini NW, editors. Sports endocrinology. Totowa (NJ): Humana Press; 2000. p. 347–70.
5. Gillespie CA, Edgerton VR. The role of testosterone in exercise-induced glycogen supercompensation. Horm Metab Res 1970;2:364–6.
6. Handelsman DJ. Indirect androgen doping by oestrogen blockade in sports. Br J Pharmacol 2008;154:598–605.
7. Donike ML, Bärwald KR, Klosterman K, et al. Nachweis von exogenem Testosteron. In: Heck H, Hollmann H, Liesen H, et al, editors. Sport: Leistung and Gesundheit. Köln (Germany): Deutscher Artze Verlag; 1982. p. 293.
8. Catlin DH, Hatton CK, Starcevic SH. Issues in detecting abuse of xenobiotic anabolic steroids and testosterone by analysis of athletes' urine. Clin Chem 1997;43:1280–8.
9. Ayotte C, Goudreault D, Charlebois A. Testing for natural and synthetic anabolic agents in human urine. J Chromatogr B Biomed Appl 1996;687:3–25.
10. Catlin DH, Cowan DA, de la Torre R, et al. Urinary testosterone (T) to epitestosterone (E) ratios by GC/MS.I. Initial comparison of uncorrected T/E in six international laboratories. J Mass Spectrom 1996;31:397–402.
11. Borts DJ, Bowers L. Direct measurement of urinary testosterone and epitestosterone conjugates using high-performance liquid chromatography/tandem mass spectrometry. J Mass Spectrom 2000;35:50–61.
12. Sottas PE, Robinson N, Saugy M, et al. A forensic approach to the interpretation of blood doping markers. Law Probab Risk 2008;7:191–210.
13. Ofterbro H. Evaluating an abnormal urinary steroid profile. Lancet 1992;339:941–2.
14. Southan G, Mallat A, Jumeau J, et al. In: Program and abstracts of the Second International Symposium on Applied Mass Spectrometry in the Health Sciences. Barcelona (Spain), April 17, 1990. p. 306.
15. Donike M, Rauth S, Mareck-Engelke U, et al. Evaluation of longitudinal studies, the determination of subject-based reference ranges of the testosteorne/epitestosterone ratio. In: Donike M, Geyer H, Gotzmann A, et al, editors. Recent advances in doping analysis, proceedings of the 11th Cologne workshop on dope analysis. Köln (Germany): Sport und Buch Strausse Edition Sport; 1994. p. 33–9.
16. Sottas PE, Saudan C, Schweizer C, et al. Bayesian detection of abnormal values in longitudinal biomarkers with an application to T/E ratio. Biostatistics 2007;8:285–96.
17. McIntosh MW, Urban N. A parametric empirical Bayes method for cancer screening using longitudinal observations of a biomarker. Biostatistics 2003;4:27–40.
18. Donike M, Mareck-Engelke U, Rauth S. Statistical evaluation of longitudinal studies, part 2: the usefulness of subject based reference ranges. In:

Donike M, Geyer H, Gotzmann A, et al, editors. Proceedings of the 12th Cologne workshop on dope analysis. Köln (Germany): Sport und Buch Strausse Edition Sport; 1995. p. 157–65.
19. World Anti-Doping Agency Technical Document TD2004EAAS. Reporting and evaluation guidance for testosterone, epitestosterone, T/E ratio and other endogenous steroids. Montreal (Canada); 2004.
20. Sottas PE, Saudan C, Schweizer C, et al. From population- to subject based limit of T/E ratio to detect testosterone abuse in elite sports. Forensic Sci Int 2008;174: 166–72.
21. Aitken C, Taroni F. Statistics and the evaluation of evidence for forensic scientists. 2nd edition. Chichester (UK): Wiley & Sons; 2004.
22. Honour JW. Urinary steroid profile analysis. Clin Chim Acta 2001;313:45–50.
23. Mareck U, Geyer H, Opfermann G, et al. Factors influencing the steroid profile in doping control analysis. J Mass Spectrom 2008;43:877–91.
24. Strahm E, Kohler I, Rudaz S, et al. Isolation and quantification by high-performance liquid chromatography-ion-trap mass spectrometry of androgen sulfoconjugates in human urine. J Chromatogr A 2008;1196:153–60.
25. Sten T, Bichlmaier I, Kuuranne T, et al. UDP-glucuronosyltransferases (UGTs) 2B7 and UGT2B17 display converse specificity in testosterone and epitestosterone glucuronidation, whereas UGT2A1 conjugates both androgens similarly. Drug Metab Dispos 2009;37:417–23.
26. Ellis L, Nyborg H. Racial/ethnical variations in testosterone levels: a probable contributor to group differences in health. Steroids 1992;57:72–5.
27. De la Torre X, Segura J, Yang Z, et al. Testosterone detection in different ethnic groups. In: Donike M, Geyer H, Gotzmann A, et al, editors. Recent advances in doping analysis (4), proceedings of the Manfred Donike workshop, 14th Cologne workshop on dope analysis. Köln (Germany): Sport und Buch Strauß; 1997. p. 71–83.
28. Santner SJ, Albertson B, Zhang G-Y, et al. Comparative rates of androgen production and metabolism in Caucasian and Chinese subjects. J Clin Endocrinol Metab 1998;83:2104–9.
29. Jakobsson J, Ekström L, Inotsum N, et al. Large differences in testosterone excretion in Korean an Swedish men are strongly associated with a UDP-Glucuronosyl transferase 2B17 polymorphism. J Clin Endocrinol Metab 2006;91(2):687–93.
30. Jakobsson-Schulze JJ, Lundmark J, Garle M, et al. Doping test results dependent on genotype of UGT2B17, the major enzyme for testosterone glucoronidation. J Clin Endocrinol Metab 2008;93:2500–6.
31. Schulze JJ, Lorentzon M, Ohlsson C, et al. Genetic aspects of epitestosterone formation and androgen disposition: influence of polymorphisms in CYP17 and UGT2B enzymes. Pharmacogenet Genomics 2008;18:477–85.
32. Xue Y, Sun D, Daly A, et al. Adaptive evolution of UGT2B17 copy-number variation. Am J Hum Genet 2008;83:337–46.
33. Strahm E, Sottas PE, Schweizer C, et al. Steroid profiles of professional soccer players: an international comparative study. Br J Sports Med 2009;43:1126–30.
34. Schulze JJ, Lundmark J, Garle M, et al. Substantial advantage of a combined Bayesian and genotyping approach in testosterone doping tests. Steroids 2009;74:365–8.
35. Falk O, Palonek E, Björkhem I. Effect of ethanol on the ratio between testosterone and epitestosterone in urine. Clin Chem 1988;34:1462–4.
36. Mareck-Engelke U, Flenker U, Schänzer W. Stability of steroid profiles. In: Schänzer W, Geyer H, Gotzmann A, et al, editors. Recent advances in doping

analysis (4), proceedings of the Manfred Donike workshop, 13rd Cologne workshop on dope analysis. Köln (Germany): Sport und Buch Strauß; 1996. p. 139–57.
37. Saugy M, Cardis C, Robinson N, et al. Test methods: anabolics. Baillieres Best Pract Res Clin Endocrinol Metab 2000;14:111–33.
38. Gelman A, Carlin JB, Stern HS, et al. Bayesian data analysis. 2nd edition. Boca Raton (FL): Chapman & Hall; 2004.
39. Sottas PE, Robinson N, Saugy M. The athlete's biological passport and indirect markers of blood doping. In: Thieme D, Hemmersbach P, editors. Doping in sports (handbook of experimental pharmacology), vol. 195. Berlin: Springer; 2010.
40. Sottas PE, Lavoué J, Bruzzi R, et al. An empirical hierarchical Bayesian unification of occupational exposure assessment methods. Stat Med 2009;28:75–93.
41. Rockman MV. Reverse engineering the genotype–phenotype map with natural genetic variation. Nature 2008;456:738–44.