

Prevalence of Blood Doping in Samples Collected from Elite Track and Field Athletes

Pierre-Edouard Sottas,^{1*} Neil Robinson,¹ Giuseppe Fischetto,² Gabriel Dollé,²
Juan Manuel Alonso,² and Martial Saugy¹

BACKGROUND: No reliable estimate of the prevalence of doping in elite sports has been published. Since 2001, the international governing body for athletics has implemented a blood-testing program to detect altered hematological profiles in the world's top-level athletes.

METHODS: A total of 7289 blood samples were collected from 2737 athletes out of and during international athletic competitions. Data were collected in parallel on each sample, including the age, sex, nationality, and birth date of the athlete; testing date; sport; venue; and instrument technology. Period prevalence of blood-doping in samples was estimated by comparing empirical cumulative distribution functions of the abnormal blood profile score computed for subpopulations with stratified reference cumulative distribution functions.

RESULTS: In addition to an expected difference between endurance and nonendurance athletes, we found nationality to be the major factor of heterogeneity. Estimates of the prevalence of blood doping ranged from 1% to 48% for subpopulations of samples and a mean of 14% for the entire study population. Extreme cases of secondary polycythemia highlighted the health risks associated with blood manipulations.

CONCLUSIONS: When applied at a population level, in this case the population of samples, hematological data can be used to estimate period prevalence of blood doping in elite sports. We found that the world's top-level athletes are not only heterogeneous in physiological and anthropometric factors but also in their doping behavior, with contrasting attitudes toward doping between countries. When applied at the individual level, the same biomarkers, as formalized in the Athlete Biological Passport paradigm, can be used in analysis of

the observed different physiological characteristics and behavioral heterogeneities.

© 2011 American Association for Clinical Chemistry

Blood doping refers to any method that aims to increase red cell mass and enhance oxygen transport capacity, thereby increasing endurance performance. Blood-doping techniques affect the endogenous production of red blood cells directly with erythropoiesis-stimulating agents (1) or indirectly with blood transfusions (2); therefore, at the end of the 1990s, some sports authorities added the measurement of markers of altered erythropoiesis to their antidoping arsenal. The introduction of full blood exams in sports combined the objectives of protecting the health of the athletes and deterring the abuse of recombinant erythropoietin (rEPO),³ which was undetectable by a drug test at that time. The International Association of Athletics Federations (IAAF), the international governing body for athletics, introduced a blood-testing program in 2001 for the world's top-level track and field athletes. Since then, the use of biomarkers to detect doping has matured into the Athlete Biological Passport (ABP) paradigm (3). As opposed to the rationale behind traditional drug tests, the fundamental principle of the ABP is that monitoring selected biomarkers over time can reveal the effects of doping on the pathology of disease. Unusually large disparities between an athlete's historic values and values acquired during a recent test may alert officials to doping or indicate a medical condition requiring closer examination (4).

In sports, many social, cultural and environmental factors affect the prevalence of doping (5), including ease of access to prohibited substances, prize money, educational programs to prevent doping, and availability and efficiency of antidoping tests. There have been several attempts to estimate the prevalence of doping in

¹ Swiss Laboratory for Doping Analyses, Epalinges, Switzerland; ² International Association of Athletics Federations, Monaco, Monaco.

* Address correspondence to this author at: Swiss Laboratory for Doping Analyses, University Centre of Legal Medicine, West Switzerland, Chemin Des Croisettes 22, 1066 Epalinges, Switzerland. E-mail Pierre-Edouard.Sottas@chuv.ch.

Received September 27, 2010; accepted February 15, 2011.

Previously published online at DOI: 10.1373/clinchem.2010.156067

³ Nonstandard abbreviations: rEPO, recombinant erythropoietin; IAAF, International Association of Athletics Federations; ABP, Athlete Biological Passport; WADA, World Anti-Doping Agency; ABPS, Abnormal Blood Profile Score; CDF, cumulative distribution function; ECDF, empirical CDF.

elite sports, mainly by use of analytical chemistry (6) or questionnaire-based surveys (7, 8). According to the statistics of the World Anti-Doping Agency (WADA), adverse or atypical analytical results occur in 1% to 2% of the tests performed in WADA-accredited laboratories (9). Although often reported in the lay sports literature, these statistics cannot reasonably be used as reliable estimates of the prevalence of doping, because drug tests based on analytical chemistry give priority to specificity at the expense of sensitivity. Imperfect sensitivity leads to false-negative results and, in the context of prevalence assessment, underestimated values. Surveys of athletes provide an interesting alternative approach, especially when methods of maintaining confidentiality, such as randomized response methods, are used to reduce individual bias in the assessment of a sensitive attribute such as doping. However, top-level athletes may be very reluctant to answer truthfully in an attempt to avoid doping suspicions directed not only toward themselves but also their sport in general.

A longitudinal record of an individual's profile of biomarkers is an invaluable tool; biomarkers of disease may assist physicians in the diagnosis of pathology, and biomarkers of doping may assist antidoping officials in its detection. On the other hand, when population biomarker data are scrutinized in the practice of evidence-based epidemiology, measures of occurrence and/or association can be derived to characterize factors affecting the health of populations. For population-based doping management, the same epidemiological methods can be used with biomarkers of doping. In particular, Bayesian inference methods have been proposed recently to model the causal relationship between blood doping and markers of an altered erythropoiesis, while taking into account heterogeneous and confounding factors (3, 10). Let us present through a simple example how the prevalence of blood doping can be estimated thanks to hematological data. The hemoglobin of a population composed of undoped Caucasian male endurance elite athletes living at low altitude is well described by a normal distribution, with a mean of 146 g/L and an SD of 9 g/L. If a blood sample is collected from 200 of these athletes, between 1 and 9 of them (4 on average) should present a value higher than 164 g/L. If 30 of these athletes presented a value higher than 164 g/L, then between 21 (11%) and 29 (15%) presented a value that was too high. Only an external cause (doping or a medical condition) can explain this discrepancy. If the prevalence of the medical conditions is known to be low, then doping is the primary cause. Although this simplistic example is biased—superior methods that use or do not use population statistics of the relevant cause (here doping) have been described elsewhere (3, 10)—it shows nevertheless that it is not necessary to have a test able to easily iden-

tify drug cheats to estimate the prevalence of test results attributable to the cause “doping”: A biomarker of doping with a known discriminative aptitude and the knowledge of the prevalence of confounding causes is enough. Today, the high standardization of the blood tests makes possible the estimation of the prevalence of blood doping by using epidemiological measures of occurrence. Here, thanks to the high number of blood tests performed since 2001, period prevalence estimates of blood doping were obtained for subgroups of samples stratified according to the athletes' sex, type of sport (endurance vs nonendurance), and nationality. Because target testing in antidoping follows a nonrandom sampling scheme, in this retrospective study we considered only the assessment of prevalence of blood doping in populations of samples. A method to estimate the prevalence of doping in populations of athletes is nonetheless proposed thereafter.

METHODS

Since 2001, a full blood count has been performed in accordance with the IAAF Blood Testing Protocol on 7289 blood samples collected from 2737 top-level track and field athletes (11). Also, data including the date of test, venue, sport, type of competition (in-, pre-, or out-of-competition), instrument (Sysmex™, Bayer-Siemens™, Coulter™, Abbott™) and date of analysis, as well as the sex, birth date, and nationality of each athlete, were collected. Because data collections were not systematic, there were missing values. As amended by the Declaration of Geneva, the nationality was anonymized to avoid any consideration of nationality from the results found in this study.

Blood samples were collected, transported, and analyzed according to the IAAF Blood-Testing Protocol (11). The multiparametric marker of doping, the Abnormal Blood Profile Score (ABPS) was calculated from the available blood profile (12) and reference cumulative distribution functions (CDFs) generated from prior knowledge acquired in clinical trials involving both controls and doped individuals (for Methods details see the Data Supplement that accompanies the online version of this article at <http://www.clinchem.org/content/vol57/issue5>). In contrast to epidemiological studies in which the prevalence of a disease is generally well established, the determination of an estimated prevalence of doping is made difficult because the doping product (e.g., rEPO, transfusion) and protocol are not known a priori. Here, a minimal estimate M1 is calculated as the maximal difference between the reference CDF that assumed no doping and the empirical CDF of the subgroup being studied. Sampling error leads to overestimated M1 for small sample sizes, but the loss in accuracy becomes negligible for large sample sizes. Sampling error aside, the measure

M1 is conservative because the size of the difference between reference and empirical CDFs depends on the sensitivity (and lack thereof) of the marker ABPS to doping. If the doping product and protocol are assumed a priori, here doping with rEPO microdoses, the measure M2 is defined as the ratio of the area between the reference CDF that assumed no doping and the empirical CDF and the area between both reference CDFs.

The 2-sample Kolmogorov–Smirnov test was used for distribution testing. For each prevalence estimate, resampling methods were implemented by constructing 1000 bootstrap estimates from the observed data set. All calculations were performed on Matlab version 7.7.0 with Statistics Toolbox version 7.0 (Mathworks).

Results

In 2001 the IAAF introduced a blood-testing program for top-level athletes. Descriptive statistics on the 7289 blood samples collected are shown in Table 1. As exemplified by the presence of 147 nationalities, the tested population was highly heterogeneous for many factors. About 5% of the athletes had been tested at an altitude higher than 2000 m. Most of these tests were performed out-of-competition, while the athletes were living or training on the high plateaus of East Africa. Similarly, of the 3-week precompetition whereabouts reported by the athletes who were tested during the 2005 and 2007 World Championships in Athletics, 4% were locations with altitudes above 2000 m. Interestingly, the altitude of the testing location followed a distribution similar to that of the altitude of the training location. In addition, endurance athletes were targeted for testing, because approximately 4 of 5 (79%) of samples collected were from athletes running distances equal to or longer than 800 m. These were athletes who could benefit from blood doping to enhance their aerobic metabolism. Finally, out-of-competition tests accounted for approximately a quarter of all the tests (23%).

Fig. 1 shows the CDFs for modal populations of female athletes for the ABPS data obtained after 2006. Five CDFs are plotted: the reference CDFs assuming no doping (left) and microdosing with rEPO (right), the empirical CDFs calculated for the samples collected from the modal group of all female athletes (green), of female athletes from country A (red), and of a subgroup including country D and other countries (blue). Other than the factors already considered, if no additional factors influenced the marker ABPS, then all empirical CDFs must fall between the 2 reference CDFs (under the assumption that the number of tests is high enough to exclude random sampling). Because of the categorization of the initial 2379 ABPS values into the subpopulations, the number of samples collected from

Table 1. Descriptive statistics on the blood samples collected from elite athletes between 2001 and 2009.

	n ^a	Occurrence
Athletes	7289	Number: 2737
Sex	7289	Male: 55%
		Female: 45%
Competition	7287	Pre: 71%
		In: 6%
		Out: 23%
Age	6266	<19 years: 6%
		19–24 years: 28%
		>24 years: 66%
Nationality	6597	A: 9.8%
		B: 7.3%
		C: 6.1%
		D: 4.7%
		E: 4.0%
		F: 3.9%
		G: 3.8%
Other: 140 <3.7%		
Ethnicity	3487	White: 63%
		Asian: 9%
		African: 27%
		Oceanian: 1%
Sport	6328	Endurance: 79%
		Nonendurance: 21%
Instrument	5584	Sysmex: 66%
		Bayer–Siemens: 26%
		Abbott: 5%
Tests altitude	3658	Coulter: 3%
		<1000 m: 93%
		1000–2000 m: 2%
Training altitude	3444	>2000 m: 5%
		<1000 m: 92%
		1000–2000 m: 4%
		>2000 m: 4%

^a Number of samples from which the statistics were derived, with the exception of “training altitude,” which represents the whereabouts of the athletes during the 3 weeks preceding the 2005 and 2007 World Championships of Athletics.

the modal group of female athletes of country A (67 samples, 53 athletes) and of the subgroup including country D (84 samples, 74 athletes) was relatively low; nevertheless, it allows for the computation of sound statistics. For example, the hypothesis that the 5 data sets of Fig. 1 are from the same distribution was re-

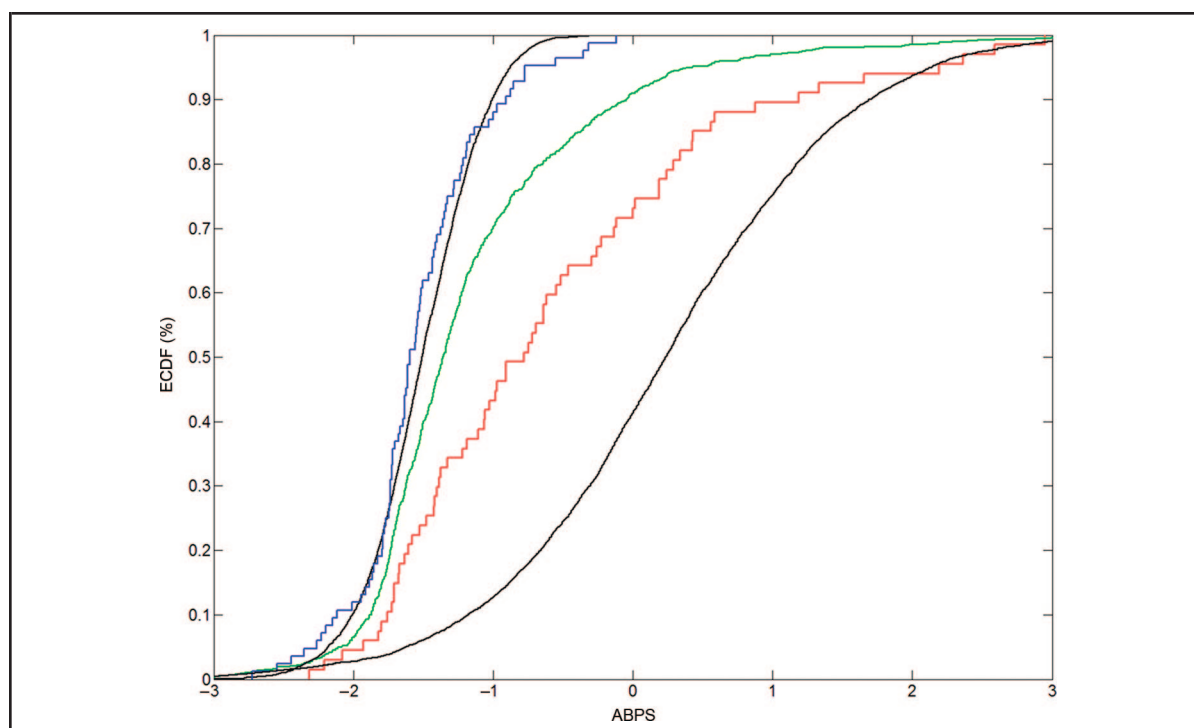


Fig. 1. Cumulative distribution functions of the multiparametric marker of blood doping ABPS based on the 7 blood variables: hemoglobin, hematocrit, red blood cell count, mean corpuscular volume, mean corpuscular hemoglobin, mean cell hemoglobin concentration, and reticulocyte percentage.

Black lines: reference CDFs obtained for a modal population of female athletes; left: assuming no doping; right: assuming doping with microdoses of rEPO (10, 13). The difference between the left and right reference CDFs represents the discriminative power of the marker ABPS (12). Other lines: empirical CDF obtained from all tests performed on all female athletes of the modal group (green, 1056 samples), on athletes of country A (red, 67 samples) and on a subgroup that includes athletes from country D (blue, 84 samples).

jected (2-sample Kolmogorov–Smirnov test: all P values <0.001), except for the comparison of the reference data that assumed no doping with the data obtained for the subgroup of country D ($P = 0.10$).

As suggested by the reference CDF that assumed no doping, it is highly unlikely that a female athlete of the modal group would present a positive abnormal blood profile score (99th percentile: -0.3). This result was confirmed by the 84 samples collected from the modal population of female athletes of country D, which returned no positive value (0 of 84 samples; 0 of 74 athletes). On the other hand, 28% of the samples (19 of 67 samples; 19 of 53 athletes) collected on the modal population of female athletes of country A returned a positive ABPS value.

Period prevalence estimates of blood doping for the various subpopulations of samples are shown in Table 2. Both minimal estimates of M1 via the difference between the reference CDF that assumed no doping and the ratio M2 between the 2 reference CDFs

assuming rEPO microdosing are shown. In addition, 95% CIs obtained by the resampling methods are shown to give insight regarding the sampling variability coming from the finite number of samples.

A snapshot of the hematological passport of 2 endurance female athletes of country A is shown in Fig. 2. Both athletes were tested according to the same protocols. The first athlete (Fig. 2A) presented variations of blood variables as expected for a healthy undoped athlete. In contrast, the second athlete (Fig. 2B) presented some unusually large variations. A closer examination was required to determine whether the polycythemia presented on 5 occasions was due to a medical condition or doping. The increased values were measured before important competitions and most probably implicated a doping behavior. Also, primary polycythemia due to an acquired or inherited genetic mutation is highly unlikely because the values returned to normal during the out-of-competition period. Independent of doping or a medical condition, health risks associated

Table 2. Period prevalence estimates of abnormal blood profiles in elite track and field athletes.

	n ^a	Prevalence M1, % ^b	Prevalence M2, %
Males	4028	12 (10–15)	12 (10–15)
Country A	205	48 (35–63)	78 (54–99)
Country B	352	3 (1–11)	1 (0–2)
Country C	257	23 (15–30)	28 (17–36)
Country D	208	6 (3–19)	5 (0–17)
Country E	160	18 (11–30)	18 (7–28)
Country F	148	6 (1–25)	2 (0–22)
Country H	160	39 (20–54)	51 (21–87)
Females	3261	18 (15–21)	18 (15–21)
Country A	445	46 (35–58)	50 (35–68)
Country B	130	8 (4–34)	2 (0–11)
Country C	147	12 (4–20)	14 (1–28)
Country D	103	1 (0–11)	0 (0–3)
Country E	106	11 (7–20)	8 (1–14)
Country F	110	6 (3–19)	0 (0–13)
Country H	65	36 (13–62)	36 (5–66)
All	7289	14 (12–16)	14 (12–16)
All nonendurance	1329	3 (0–8)	1 (0–3)
All endurance	4999	18 (15–22)	19 (16–22)

^a n, Number of samples from which the estimates were derived. Prevalence M1, minimal estimates without any assumptions on the doping method. Prevalence M2, estimates obtained assuming doping with rEPO microdoses.

^b 95% CI estimated by bootstrapping methods, with any negative estimates rounded toward 0%.

with polycythemia include a higher risk of thrombus and clot formation that may lead to strokes, heart attack, and pulmonary embolism. This risk may be particularly problematic for endurance athletes such as long distance runners because they are more subject to the loss of body fluid and hypovolemia.

Discussion

Depending on the population, the period prevalence estimates, M1, ranged from 1% to approximately 48%. Because all athletes participated in the same competitions or underwent out-of-competition testing according to the same protocols, any confounding factors related to the procedure or analysis were safely excluded. In particular, any systematic error from a lack of calibration was unlikely for period prevalence estimates. Rather than for period prevalence estimates, a good calibration is particularly important in point prevalence estimates to avoid any systematic bias, for exam-

ple, when all athletes participating in the same competition are tested at the same time.

Although strict standardization was applied for the collection, transport, and analysis of the blood samples, standardization of methods can be checked from the data itself. In particular, because ABPS uses volumetric blood variables such as mean corpuscular volume, the marker is also sensitive to any deterioration in the blood samples that may have occurred during the transport of the samples. Fig. 1 shows the ABPS results for 2 subgroups tested according to the same protocols as athletes participating in the same competitions. The fact that some subgroups [e.g., athletes from country D (Fig. 1)] produced an empirical CDF (ECDF) very close to the CDFs used as a reference confirms that the standardization was rigorous enough for drawing sound conclusions. In particular, only an external effect can explain the difference in the ECDFs between countries A and D in Fig. 1, because both groups were tested according to the same protocols. In Fig. 1, if the ECDF obtained from athletes of country D is used as the reference instead of the generated CDF, then the measure M2 is equal to 43% (27%–60%), suggesting that the prevalence of doping is about 43% higher in country A than in country D. This value is in good agreement with the difference (46% – 1% = 45%) found when all samples are taken from country A and D, i.e., when the subgroup is not limited to ABPS data after 2006 and to the modal group. All these considerations suggest that the validity of the period prevalence estimates M1 and M2 are not affected by a lack of standardization in blood data acquisition.

The prevalence of some blood disorders can be relatively high in some regions, such as thalassemia among Mediterranean people and sickle cell anemia among Sub-Saharan African people. Here, the high values of the measures M1 and M2 cannot be attributed to these disorders, because the biomarker of doping ABPS returns abnormally low—instead of abnormally high—values for blood profiles measured from individuals with thalassemia or anemia. Other rare hematological anomalies, such as myeloproliferative diseases and Chuvash polycythemia, can on the other hand cause high ABPS values at the individual level. However, at the population level, their prevalence is known to be significantly lower [e.g., 22 per 100 000 for polycythemia vera and 24 per 100 000 for essential thrombocytopenia in the US (13)] than some of the measures estimated here. Therefore, at the population level, secondary, artificially induced polycythemia remains the main cause of the high estimates found for some groups of the world's top-level athletes.

In this retrospective study, the measures M1 and M2 represent prevalence estimates in populations of samples, not in populations of individuals. First, to op-

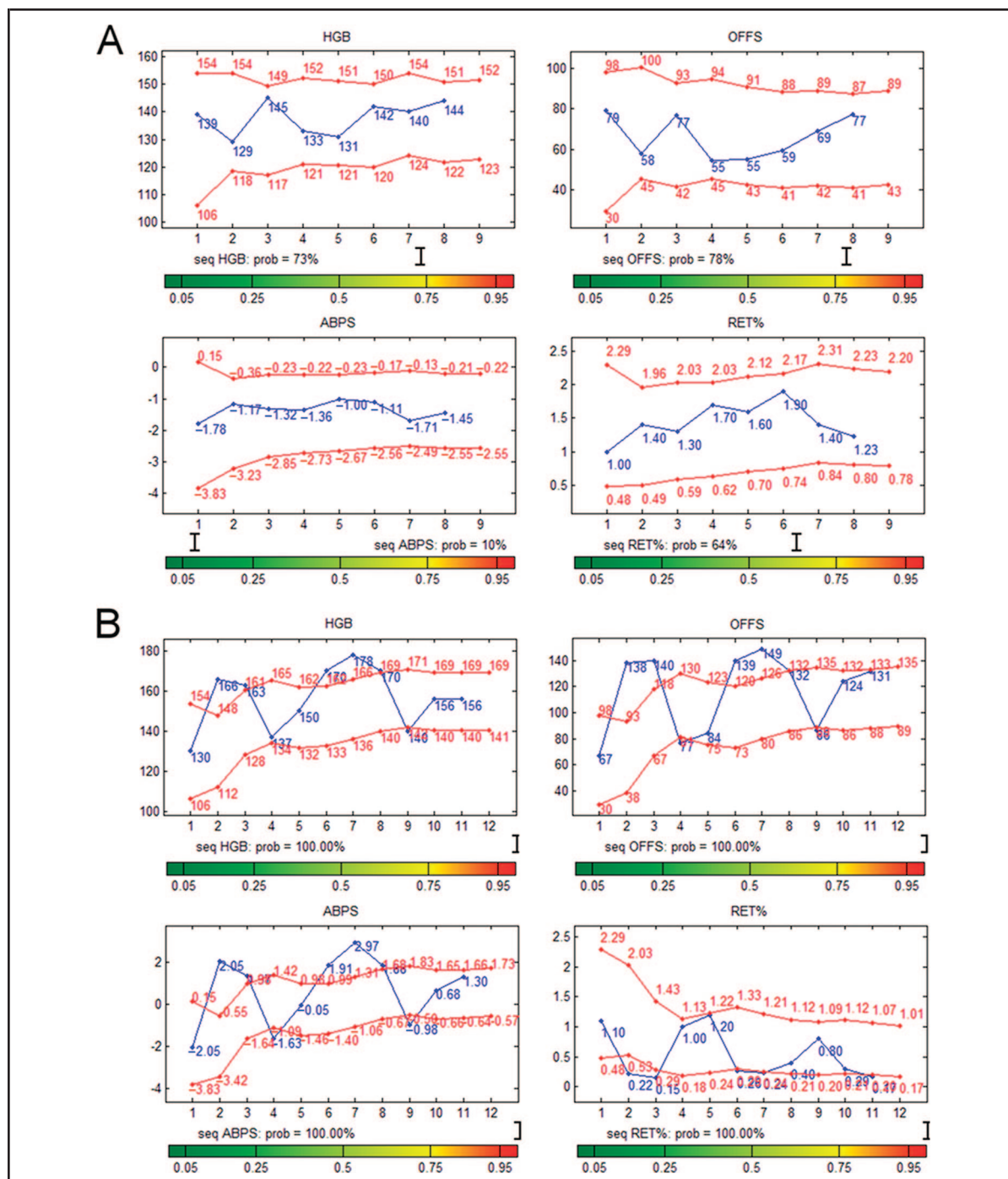


Fig. 2. Snapshots of the hematological module of the ABP of 2 (A, B) Caucasian female endurance athletes from country A.

Upper left, hemoglobin (HGB); upper right: stimulation index OFF-score (OFFS); lower left: ABPS; lower right: reticulocyte percentage (RET%). Blue lines: actual test results (8 tests between 2003 and 2006 for the first athlete, 11 tests between 2005 and 2008 for the second), red lines: individual limits as found for the Athlete Hematological Passport [Sottas et al. (3)]. The colored bars indicate at which percentile of the distribution of expected sequences falls the observed sequence of values; seq, sequence; prob, probability.

timize the deterrent effect of blood testing and the cost-effectiveness of the program, the test distribution plan privileged abnormal values: an athlete presenting an abnormal blood profile, whether it was due to a medical condition or doping, was tested more often than an athlete who presented normal values in the first test. Nonrandom sampling among athletes makes the prevalence in populations of samples higher than the prevalence in populations of individuals. Similarly, blood samples were also collected more frequently from endurance athletes. Unsurprisingly, we found a significant difference in the prevalence estimates according to this factor [M1: (0%–8%) for nonendurance, (15%–22%) for endurance]. Any prevalence estimate found on a subpopulation of samples stratified according to factors such as sex, type of sport, and country cannot be extended to the full population of track and field athletes. Second, the period prevalence estimates are independent of the doping protocol and, more particularly, its duration. For health (and cost) reasons, it is very unlikely that an unscrupulous athlete would abuse a doping product, such as rEPO, 365 days a year. In the second passport depicted in Fig. 2, for example, at least 4 tests (tests 1, 4, 5, and 9) are representative of a normal, undoped athlete.

As shown in Table 2, the number of samples was high enough to obtain reasonably precise estimates of the prevalence of doping in samples. As a result, it has been possible to point out that the prevalence of blood doping is highly dependent on the athlete's country of origin, with significantly different estimates between countries. The strong correlation between the prevalence estimates of males and females as well as higher estimates for the endurance athletes gives confidence to the developed method and, in turn, to the detected disparities between countries. In addition to an inefficient antidoping testing policy, a poor prevention policy may help to explain the presence of a doping culture in some countries.

The methods presented here for the estimation of the prevalence of doping were applied to populations of samples on a retrospective basis. Interestingly, there is no conceptual restriction to the application of the same methods to a series of values obtained from the same individual that are applied to populations of sequences. From a Bayesian inference perspective, in the present study we focused on the comparison of nonrandomly selected data with the prior predictive distributions of ABPS values in factor-stratified subpopulations of samples. Interestingly, the necessary mathematical framework has already been developed to use the same Bayesian network to compare sequences of biomarkers of doping with the prior predictive distribution of sequences in factor-stratified subpopulations of athletes (3). The prospective application of the latter method would return estimates of blood dop-

ing in populations of athletes. The main requirement is the implementation of a longitudinal hematological passport according to the ABP paradigm (3, 4). Finally, any selection bias can be either overcome with a dedicated test distribution plan or at least corrected by a statistical method such as the Heckman correction (14).

The IAAF introduced blood tests in 2001 for possibly the most heterogeneous population ever tested. A strong push to follow the protocol started in 2005; since then, no fewer than 1500 blood samples have been collected yearly. Of course, the multiethnic characteristics of track and field athletes make it difficult to practically implement health rules for competition on the basis of only a single test evaluation. On the other hand, the use of decision support systems that are scientifically validated and recognized may be an effective strategy to achieve 2 main results simultaneously. First, this process develops a clear knowledge source of the different individual athletes' particular athlete populations, which may differ according to not only ethnic, physiologic, and environmental characteristics but also other, sometimes artificially induced, external factors. Second, within each ethnically homogeneous population, it is necessary to clearly highlight the abnormal profiles (of medical or nonmedical origin) and to target individual athletes with further tests, including urine and/or blood tests. Following the same concept as personalized medicine, an individual hematological passport can provide, according to international rules and after exclusion of medical and physiological conditions, strong evidence of blood manipulation that is formally and legally accepted. As a result, the implementation of the ABP paradigm in elite track and field athletics may not only level the playing field but also protect the athletes' health.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors' Disclosures or Potential Conflicts of Interest: Upon manuscript submission, all authors completed the Disclosures of Potential Conflict of Interest form. Potential conflicts of interest:

Employment or Leadership: None declared.

Consultant or Advisory Role: None declared.

Stock Ownership: None declared.

Honoraria: None declared.

Research Funding: P.-E. Sottas, World Anti-Doping Agency grant no. R07D0MS.

Expert Testimony: None declared.

Role of Sponsor: The funding organizations played no role in the design of study, choice of enrolled patients, review and interpretation of data, or preparation or approval of manuscript.

References

1. Macdougall IC, Ashenden M. Current and upcoming erythropoiesis-stimulating agents, iron products, and other novel anaemia medications. *Adv Chronic Kidney Dis* 2009;16:117–30.
2. Giraud S, Sottas PE, Robinson N, Saugy M. Blood transfusion in sports. In: Thieme D, Hemmersbach P, eds. *Doping in sports: biochemical principles, effects and analysis. Handbook of experimental pharmacology* 195. Berlin: Springer; 2010.
3. 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.
4. Ashenden MJ. A strategy to deter blood doping in sport. *Haematologica* 2002;87:225–32.
5. Sjöqvist F, Garle M, Rane A. Use of doping agents, particularly anabolic steroids, in sports and society. *Lancet* 2008;371:1872–82.
6. Thevis M, Sauer M, Geyer H, Sigmund G, Mareck U, Schänzer W. Determination of the prevalence of anabolic steroids, stimulants, and selected drugs subject to doping controls among elite sport students using analytical chemistry. *J Sports Sci* 2008;26:1059–65.
7. Scarpino V, Arrigo A, Benzi G, et al. Evaluation of prevalence of doping among Italian athletes. *Lancet* 1990;336:1048–50.
8. Striegel H, Ulrich R, Simon P. Randomized response estimates for doping and illicit drug use in elite athletes. *Drug Alcohol Depend* 2010;106:230–2.
9. World Anti-Doping Agency. <http://www.wada-ama.org/en/Anti-Doping-Community/Anti-Doping-Laboratories/Laboratory-Statistics/> (Accessed March 2010).
10. Sottas PE, Robinson N, Saugy M, et al. A forensic approach to the interpretation of blood doping markers. *Law Prob Risk* 2008;7:191–210.
11. International Association of Athletics Federations. Rules and regulations. <http://www.iaaf.org/antidoping/rules/other/index.html> (Accessed March 2010).
12. Sottas PE, Robinson N, Giraud S, et al. Statistical classification of abnormal blood profiles in athletes. *Int J Biostatistics* 2006;2:3.
13. Ma X, Vanasse G, Cartmel B, Wang Y, Selinger HA. Prevalence of polycythemia vera and essential thrombocythemia. *Am J Hematol* 2008;83:359–62.
14. Heckman J. Sample selection bias as a specification error. *Econometrica* 1979;47:153–61.