REVIEW ARTICLE

Sports Med 2009; 39 (2): 107-127
0112-1642/09/0002-0107/$49.95/0

Sea-Level Exercise Performance Following Adaptation to Hypoxia

A Meta-Analysis

Darrell L. Bonetti and *Will G. Hopkins*

Institute of Sport and Recreation Research, AUT University, Auckland, New Zealand

Contents

Abstract

Adaptation to living or training in hypoxic environments (altitude training) continues to gain interest from sport scientists and endurance athletes. Here we present the first meta-analytic review of the effects on performance and related physiological measures following adaptation to six protocols of natural or artificial hypoxia: live-high train-high (LHTH), live-high train-low (LHTL), artificial LHTL with daily exposure to long (8–18 hours) continuous, brief (1.5–5 hours) continuous or brief (<1.5 hours) intermittent periods of hypoxia, and artificial live-low train-high (LLTH).

The 51 qualifying studies provided 11–33 estimates for effects on power output with each protocol and up to 20 estimates for effects on maximal oxygen uptake ($\dot{V}O_{2max}$) and other potential mediators. The meta-analytic random-effect models included covariates to adjust for and estimate moderating effects of study characteristics such as altitude level and days of exposure. Poor reporting of inferential statistics limited the weighting factor in the models to sample size. Probabilistic inferences were derived using a smallest worthwhile effect on performance of 1%. Substantial enhancement of maximal endurance power output in controlled studies of subelite athletes was very likely with artificial brief intermittent LHTL (2.6%; 90% confidence limits ±1.2%), likely with LHTL (4.2%; ±2.9%), possible with artificial long continuous LHTL (1.4; ±2.0%), but unclear with LHTH (0.9; ±3.4%), artificial brief continuous LHTL (0.7%; ±2.5%) and LLTH (0.9%; ±2.4%). In elite athletes, enhancement was possible with natural LHTL (4.0%; ±3.7%), but

unclear with other protocols. There was evidence that these effects were mediated at least partly by substantial placebo, nocebo and training-camp effects with some protocols. Enhancing protocols by appropriate manipulation of study characteristics produced clear effects with all protocols (3.5–6.8%) in subelite athletes, but only with LHTH (5.2%) and LHTL (4.3%) in elite athletes. For $\dot{V}O_{2max}$, increases were very likely with LHTH (4.3%; ±2.6%) in subelite athletes, whereas in elite athletes a 'reduction' was possible with LHTH (−1.5%; ±2.0%); changes with other protocols were unclear. Effects on erythropoietic and other physiological mediators provided little additional insight into mechanisms.

In summary, natural LHTL currently provides the best protocol for enhancing endurance performance in elite and subelite athletes, while some artificial protocols are effective in subelite athletes. Likely mediators include $\dot{V}O_{2max}$ and the placebo, nocebo and training-camp effects. Modification of the protocols presents the possibility of further enhancements, which should be the focus of future research.

When an athlete ascends from sea level to moderate altitude, the shortage of oxygen (hypoxia) initially impairs endurance training and performance. After a few weeks at altitude, training and performance recover to some extent as the athlete adapts. If the athlete then returns to sea level, do the adaptations lead to enhancement of endurance performance? Coaches have long thought so, but studies aimed at this question appeared to be inconclusive, leading researchers to suspect that any benefit from adaptation to hypoxia was offset by loss of endurance fitness consequent to the reduction in training intensity.[1] The focus of research on altitude training then moved from this traditional 'live-high train-high' approach (LHTH) to live-high train-low (LHTL), in which athletes live and sleep at altitude, but descend regularly to lower altitude for training sessions.[2] LHTL appeared to be more successful, and interest has grown in the use of nitrogen houses, hypobaric chambers, altitude tents or hypoxic inhalers to adapt to hypoxia and train normally without having to travel up and down a mountain.[1] Researchers have investigated three such approaches to artificial LHTL: (i) continuous exposure to a simulated moderate altitude for periods of 8–18 hours per day (artificial long continuous LHTL); (ii) continuous exposure to a simulated moderate-high altitude for 1.5–5 hours per day (artificial brief continuous LHTL); and (iii) intermittent exposure to a simulated high altitude for <1.5 hours per day (artificial brief intermittent LHTL). The same devices have also been used to simulate moderate altitude while the athlete exercises continuously or intermittently for at least 0.5 hours per session (artificial LLTH).

While there is general agreement that adaptation to some forms of hypoxia can enhance sea-level performance, there has been considerable debate recently about the physiological mechanisms.[3-5] Gore and Hopkins[4] provided a rationale for understanding the mechanisms underlying the effects on maximal performance of differing durations. Exercise intensities below maximal oxygen uptake ($\dot{V}O_{2max}$) [>10 minutes' duration] are sustained essentially by aerobic power, whereas exercise intensities above $\dot{V}O_{2max}$ are sustained by a combination of aerobic and anaerobic power. Aerobic power consists of three components: (i) $\dot{V}O_{2max}$; (ii) the fraction of maximal uptake that can be sustained during the exercise; and (iii) economy or efficiency of conversion of oxygen consumption into power output.[6] Changes in endurance performance following adaptation to hypoxia could therefore be due to changes in any of these three components, along with any changes in the contribution of anaerobic power for supramaximal exercise. Researchers who are interested in the mechanisms

underlying the performance effects of hypoxic adaptation measure one or more of these components or the more fundamental physiological variables underlying them.

There has been no previous meta-analytic review of the effects on performance or related physiological measures following adaptation to any of the artificial or natural forms of altitude training. The current review addresses this deficit. There were sufficient studies to allow us to meta-analyse separately the effects on performance of the six natural and artificial altitude protocols. By far the most popular potential mechanism variable has been $\dot{V}O_{2max}$, and we have also been able to meta-analyse this variable with all six protocols. Researchers have long argued that enhancements in $\dot{V}O_{2max}$ are mediated by erythropoiesis,[1,5,7-9] so measurements of erythropoietin, reticulocytes, red cell mass, haemoglobin mass, haemoglobin concentration and ferritin have also been reported. We have been able to meta-analyse haemoglobin concentration for LHTH and artificial brief intermittent LHTL, but we had to meta-analyse haemoglobin mass and red cell mass by combining them across all protocols. We were able to perform only a graphical analysis for erythropoietin, reticulocytes and ferritin due to the small number of estimates for these variables. Mechanisms underlying anaerobic power are less popular with researchers, and only the indirect measure of anaerobic power represented by peak blood lactate following an exercise test was reported in sufficient studies for meta-analysis in LHTH and artificial brief intermittent LHTL.

1. Methodology

1.1 Study Selection

Searches of PubMed, SportDiscus and Google Scholar were performed for studies published in English up to and including April 2007. Reference lists in review and original research articles identified were also examined. The primary focus of the meta-analysis was performance. We therefore included studies of performance measured at or near sea level (<1000 m). Studies published only as conference abstracts were not excluded. We included studies with measures of oxygen consumption directly related to endurance performance, but studies reporting haematological or other parameters not directly related to performance and lacking a performance measure were excluded. Several studies were excluded for poor reporting of data or for not assessing performance at or near sea level.[10-25] Other reasons for excluding studies were: a performance enhancement of 19% in 5 mmol/L lactate speed in elite runners, when other measures of performance increased by 0.6% and 1.1%;[26] the only uncontrolled study in LLTH and with only five athletes;[27] and poor compliance with training, a non-specific performance test, and an uncertain post-exposure test time in an uncontrolled study of the brief continuous LHTL protocol.[28] The descriptive statistics for the 51 qualifying studies are shown in table I.

1.2 Data Extraction

The study estimates for the treatment effect were calculated for estimates without a control group by dividing the mean post-score by the mean pre-score for the experimental group and expressing the ratio as a percentage; for estimates with a control group, the post-/pre-score ratio in the experimental group was divided by the post-/pre-score ratio in the control group before converting to a percentage. Percentage change in performance time in time trials was converted to change in mean power output by multiplying by an appropriate factor derived from power-velocity relationships.[73] For running, the factor was −1; for cycling, the factor was −2.5; for swimming, the factor was −2.0, which was an index x derived from first principles[73] by fitting the power-velocity relationship $P = kV^x$ to published data.[74] For any exercise modality, the percentage change in time to exhaustion at a constant power was converted to percentage change in power output in an equivalent time trial of the same duration by multiplying by a factor derived from models for the power-duration relationship of human performance, as follows: for

Table I. Characteristics of study groups included in the meta-analysis sorted by protocol and first author

Study	Subjects	Sample size[a]	Design	Competitive level	Training phase	Hypoxic (h/d)[b]	Exposure/ intervention days[c]	Altitude level (m)[d]	Hypoxia device
Live-high train-high									
Bailey et al.[29]	Runners	8M, 2F; 14M, 5F	C	Elite	?	24	28	1640	
	Runners	9M, 5F; 6M, 3F	C	Elite	?	24	28	1750	
Burtscher et al.[30]	Runners	10M; 12M	C	Subelite	?	24	12	2315	
Friedmann et al.[31]	Boxers +Fe[e]	9M	U	Subelite	Off-season	24	18	1800	
	Boxers −Fe[e]	7M	U	Subelite	Off-season	24	18	1800	
Gore et al.[32]	Cyclists	8M	U	Elite	?	24	31	2690	
Ingjer and Myhre[33]	Skiers	7M; 7M	U	Elite	Competitive	24	21	1900	
Jensen et al.[34]	Rowers	9M; 9M	C	Elite	Competitive	24	21	1822	
Levine and Stray-Gundersen[35]	Runners	10?	U	Subelite	?	24	28	1200	
	Runners	9?	U	Subelite	?	24	28	2500	
Levine and Stray-Gundersen[2]	Runners	9M, 4F; 9M, 4F	C	Subelite	Competitive	24	28	2500	
Miyashita et al.[36]	Swimmers	12M, 8F	U	Elite	Competitive	24	21	2300	
Pyne[37]	Swimmers	14M, 8F	U	Elite	Competitive	24	21	2102	
Rusko et al.[38]	Skiers	14M; 7M	C	Elite	?	24	22	1700	
Saunders et al.[39]	Runners	10M; 13M	C	Elite	?	24	20	1750	
Svedenhag and Saltinj[40]	Runners	5M; 4M, 2F	C	Elite	?	24	14	2000	
Svedenhag et al.[41]	Skiers	5M, 2F	U	Elite	?	24	30	1900	
Live-high train-low									
Dehnert et al.[42]	Triathletes	6?; 10?	C	Subelite	?	~18–24	13	1956/800	
Levine and Stray-Gundersen[2]	Runners	9M, 4F; 9M, 4F	C	Subelite	Competitive	~18–24	28	2500/1200	
Stray-Gundersen and Levine[43]	Runners	6?	U	Subelite	?	~18–24	28	2500/1200	
Stray-Gundersen et al.[8]	Runners	8F, 14M	U	Elite	Competitive	~18–24	27	2500/1200	
Wehrlin et al.[44]	Orienteers	5M, 5F	U	Elite	Pre-season	~18–24	24	2456/1000	
Witkowski et al.[45]	Runners	8M, 4F	U	Subelite	?	~18–24	28	1780/1250	
	Runners	8M, 4F	U	Subelite	?	~18–24	28	2085/1250	
	Runners	8M, 4F	U	Subelite	?	~18–24	28	2454/1250	
	Runners	8M, 4F	U	Subelite	?	~18–24	28	2805/1250	
Artificial long continuous live-high train-low									
Clark et al.[46]	Cyclists, triathletes	9M; 10M	C	Subelite	?	9–10	20	2650	N_2 house
	Cyclists, triathletes	10M; 10M	C	Subelite	?	9–10	20/24	2650	N_2 house

Continued next page

Table I. Contd

Study	Subjects	Sample size[a]	Design	Competitive level	Training phase	Hypoxic (h/d)[b]	Exposure/ intervention days[c]	Altitude level (m)[d]	Hypoxia device
Gore et al.;[47] Hahn et al.[48]	Triathletes	6M; 6M	C	Elite	?	8–10	23	3000	N_2 house
Hahn et al.[48]	Cyclists	5F; 6F	C	Elite	?	8–10	12	2650	N_2 house
Hinckson and Hopkins[49]	Runners, triathletes	11M; 11M	X	Subelite	?	8	25	2500–3500	N_2 tent
Hinckson et al.[50]	Runners	8M, 2F	U	Subelite	?	10	24–30/30	2500–3500	N_2 tent
Martin et al.[51]	Cyclists	5F; 6F	C	Elite	?	8–10	12	2650	N_2 house
Mattila and Rusko[52]	Cyclists	5M	U	Elite	Competitive	18	11	3000	N_2 house
Nummela[26]	Runners	6M, 2F; 10M	C	Elite	?	16–17	17	2200	N_2 house
Roberts et al.[53]	Cyclists	14M, 5F; 14M, 5F	X	Subelite	?	8–10	5–15	2650	N_2 house
Rusko et al.[38]	Skiers	9M, 3F; 8M, 2F	B?	?	?	12–16	25	2500	N_2 house
Saunders et al.[39]	Runners	10M; 13M	C	Elite	?	9–12	19/25	2000–3100	N_2 house
Artificial short continuous live-high train-low									
Basset et al.[54]	Skiers, skaters	7M, 5F; 7M, 5F	X, B	Subelite	Off-season	3	6/19	3650	N_2 tent
Katayama et al.[55]	Runners	6M; 6M	C	Subelite	?	1.5	9/19	4000	Chamber
Katayama et al.[56]	Runners	8M; 7M	C	Subelite	Competitive	3	14	4000	N_2 tent
Gore (2006); Rodriguez (2007)[57,58]	Swimmers	3M, 3F; 4M, 3F	C	Subelite	?	3–5	9	4000–5500	Chamber
Gore et al.;[57] Rodriguez et al.[58]	Runners	2M, 3F; 3M, 2F	C	Subelite	?	3–5	9	4000–5500	Chamber
Artificial brief intermittent live-high train-low[f]									
Bonetti et al.[59]	Kayakers	10M; 10M	X	Subelite	Competitive	0.5/1	15/19	3600–6000	Inhaler
Bonetti et al.[60]	Cyclists	18M; 9M	C	Subelite	Competitive	0.5/1	15/19	3600–6000	Inhaler
Hamlin and Hellmans[61]	Multisport athletes	5M, 7F; 8M, 2F	C, B	Subelite	?	0.75/1.5	15/19	3400–5000	Inhaler
Hinckson et al.[62]	Rowers	2M, 5F; 1M, 4F	C, B	Elite	?	0.9/1.5	15/19	3600–6000	Inhaler
Julian et al.[63]	Runners	7M; 7M	C, B	Elite	Competitive	0.75/1.5	20/26	3600–5000	Inhaler
Wood et al.[9]	Hockey players	15M; 14M	C, B	Subelite	Competitive	0.6/1	15/19	3600–6000	Inhaler
Live-low train-high									
Dufour et al.[64]	Runners	9M; 9M	C	Subelite	Pre-season	0.2–0.33/0.33	12/40	3000	Inhaler
Hendriksen and Meeuwsen[65]	Triathletes	12M; 12M	X, B?	Subelite	Pre-season	2	10	2500	Chamber
Katayama et al.[66]	Non-athletes	7M; 7M	C	Trained	?	0.5	10/12	4500	Chamber
Morton and Cable[67]	Team sports	8M; 8M	C	Trained	?	0.17/0.5	9/19	2750	N_2 house

Continued next page

Table I. Contd

Study	Subjects	Sample size[a]	Design	Competitive level	Training phase	Hypoxic (h/d)[b]	Exposure/ intervention days[c]	Altitude level (m)[d]	Hypoxia device
Roels et al.[68]	Cyclists, triathletes	11M; 11M	C, B?	Subelite	Pre-season	0.2–0.5/0.5	14/47	3000	Inhaler
Terrados et al.[69]	Cyclists	4M; 4M	C	Subelite	?	2	20/20–26	2300	Chamber
Truijens et al.[70]	Swimmers	3M, 5F; 3M, 5F	C, B	Subelite	?	0.21/0.5	15/33	2500	Inhaler
Ventura et al.[71]	Cyclists	6M; 5M, 1F	C, B?	Subelite	Competitive	0.5	18/40	3200	Inhaler

a Data separated by ';' are controlled trials with sample size in experimental and control groups.

b Numbers separated by '/' indicate sum of time in bouts of hypoxia and sum of recovery time per session.

c Numbers after '/' indicate intervention period, if longer than exposure period.

d Numbers separated by '/' indicate live-high and train-low altitudes.

e Groups with and without iron supplementation.

f Altitude level estimated from arterial oxygen saturation in each study using the figure at www.high-altitude-medicine.com/SaO2-table.html.[72]

B=blind; **B?** indicates blinding uncertain (assumed not blind); **C**=controlled trial; **F**=female; **M**=male; **U**=uncontrolled trial; **X**=crossover; **?** indicates uncertain.

supramaximal tests (<7.5 minutes), the factor was 0.50/T, where T is the time in minutes;[75] for submaximal tests (>7.5 minutes), the factor was approximately 1/15.[73] Percentage change in time to exhaustion in incremental tests was converted to percentage change in peak power by multiplying by a factor 1-f, where f was the power of the first stage of the test expressed as a fraction of the peak power, under the assumption that the load increased linearly to maximum. A spreadsheet containing all study estimates can be obtained from the authors.

1.3 Meta-Analyses

The main outcome from a meta-analysis is a weighted mean of values of an outcome statistic from various studies, where the weighting factor is usually the inverse of the square of the sampling standard error of the statistic. The standard error is derived from either the confidence interval or p-value of the statistic or from standard deviations of change scores in control and experimental groups. Unfortunately, 55% of the study-estimates for performance that would have otherwise qualified for inclusion in our meta-analyses did not have sufficient information to derive the standard error; for estimates from studies other than of intermittent artificial LHTL, the figure is 71%. The main problem was reporting of statistical significance or non-significance as a p-value inequality without any further inferential information. To exclude all these studies from the meta-analyses would have resulted in unacceptable bias, akin to the publication bias that arises from failure of authors to submit studies with non-significant outcomes or failure of journal editors to accept them. We therefore performed the meta-analyses with a weighting factor derived from the sample size for each study estimate. The factor was (study sample size)/(mean study sample size). To calculate the sample size equivalent to that of a parallel-groups controlled trial with equal-sized groups when the groups were of unequal size n_1 and n_2, we assumed equal standard error e in each group. The standard error of the difference in means between groups is therefore $e^2/n_1 + e^2/n_2$, and for

groups of equal size n the standard error is $2e^2/n$. It follows that the effective sample size = $2n = 4n_1n_2/(n_1+n_2)$. An uncontrolled trial is equivalent to a controlled trial in which the control group has a mean of zero and an infinite sample size, but as $n_1 \rightarrow \infty$, $4n_1n_2/(n_1+n_2) \rightarrow 4n_2$, so the sample size for uncontrolled trials was inflated by a factor of 4 to make it equivalent to that of a controlled trial. To ensure studies with different numbers of estimates would have equal weighting, each study's weighting factor was divided by the number of estimates it provided and multiplied by the mean number of estimates in all the studies contributing to the meta-analysis. The resulting meta-analysed effect is equivalent to that produced in a random-effect meta-analysis in which the between-study variance far outweighs the error variance in each study estimate, so the confidence interval must be more conservative (wider) than would be provided by the usual random-effect analysis. An assumption underlying our analyses is that the dependent variable giving rise to the study estimates has the same error of measurement in all studies, but violation of this assumption will result only in minor differences in the weight given to each study; the main differences in weight arise from differences in effective sample size.

The meta-analyses were performed with the mixed modelling procedure (Proc Mixed) in the Statistical Analysis System (Version 9.2, SAS Institute, Cary, NC, US). Percentage effects were converted to factors (= 1 + effect/100), log transformed for the analysis, then back transformed to percentages. Study characteristics were the fixed effects in the model; these were included as main effects only because of the limited number of study estimates. We limited the characteristics to those that were included in most studies and that might be expected on physiological or psychological grounds to moderate the effect of hypoxia: competitive status (elite vs subelite); design characteristics (uncontrolled vs controlled trial, non-blind vs blind trial); sex (males as a fraction of the sample); training phase (competitive vs non-competitive or unknown); altitude level or its equivalent for artificial altitude (m); hours of hypoxia per day (for LHTH, LHTL and artificial long-duration LHTL); minutes of hypoxia per day (for the remaining protocols, not counting minutes spent in normoxia between intervals of hypoxia); count of days when any exposure to hypoxia occurred; total count of treatment days, including any days resting from exposure; ratio of exposure/treatment days; day post-exposure when performance was tested; training intensity on a 1–4 scale (for LLTH); type of performance test (submaximal vs maximal); and duration of maximal exercise tests (minutes). Missing values for sex of nine and ten subelite runners experiencing LHTH[35] and of six subelite runners experiencing LHTL[43] were assigned the mean value of proportion of males for their protocols. Competitive status was deemed elite if the athletes were in a national team and competing at international level. The four points of the training-intensity scale for LLTH were: above $\dot{V}O_{2max}$, 4; around $\dot{V}O_{2max}$, 3; around anaerobic threshold, 2; below anaerobic threshold, 1. Post-exposure test day and duration of maximum exercise tests were log transformed before analysis and included as simple linear predictors. Supplementary analyses (not shown in table II) were also performed, where possible, with post-exposure test day included as a quadratic or cubic polynomial in ×/÷ standard deviation units, to investigate the possibility of peaks or troughs in performance.

An effect of a study characteristic is not shown in the tables for one or more of the following reasons: there was insufficient variation in the characteristic between-study estimates to estimate the effect; collinearity with other study characteristics prevented its estimation; and the small number of study estimates limited the analysis to only a few characteristics. To compare effectiveness of protocols on performance, the meta-analysed effects are shown for subelite athletes (all protocols) and elite athletes (four protocols) and are adjusted to 100% controlled trials and 100% maximal tests. For all the other study characteristics, we could not adjust to the same common value, so the effects on performance for each protocol are shown evaluated at the mean values of the study characteristics for that protocol.

Table II. Meta-analysis of effects on sea-level mean power output following adaptation to hypoxia experienced in studies with various protocols of natural and artificial altitude. Effects of mean and enhanced protocols are those predicted for controlled trials and maximal tests. Effects in parentheses are unclear (>5% chance of enhancement and >5% chance of impairment); otherwise **bold** indicates ≥50% chance of enhancement, *italic* indicates ≥50% chance of impairment, and plain font indicates ≥50% chance of trivial effect. These probabilistic outcomes are computed with reference to a smallest important change of 1%

Effect	Natural altitude protocols		Artificial altitude protocols			
	live-high train-high	live-high train-low	live-high 8–18 h/d, continuous, train-low	live-high 1.5–5 h/d, continuous, train-low	live-high <1.5 h/d, intermittent, train-low	live-low train-high 0.5–2 h/d
Effect of mean protocol[a] (%); ±90% CL[b]						
Elite	(1.6; ±2.7)	**4.0; ±3.7**	(0.6; ±2.0)		(0.2; ±1.8)	
Subelite	(0.9; ±3.4)	**4.2; ±2.9**	**1.4; ±2.0**	(0.7; ±2.5)	**2.6; ±1.2**	(0.9; ±2.4)
Effect of enhanced protocol[c] (%); ±90% CL						
Elite	**5.2; ±4.1**	**4.3; ±4.1**	(4.0; ±5.5)		(1.2; ±2.5)	
Subelite	**4.5; ±4.1**	**4.6; ±3.3**	**4.8; ±5.3**	**3.5; ±3.5**	**3.6; ±2.1**	**6.8; ±4.9**
Study characteristics changed by +1 SD or −1 SD for enhanced protocol	+ Altitude − Days exposure + Test day	− Altitude − Test day	+ Altitude + Hours hypoxia − Days exposure	− Altitude − Test day	+ Exposure ratio − Test day	− Altitude − Train intensity + Days exposure + Test day
Study characteristics (mean ± SD)						
References	10	5	9	4	6	7
Study groups	13	9	10	5	5	7
Study estimates	33	13	17	11	33	19
Subjects/estimate	16 ± 7	12 ± 6	17 ± 9	15 ± 5	20 ± 6	17 ± 5
Effective subjects/estimate	36 ± 22	41 ± 11	20 ± 9	15 ± 5	20 ± 6	17 ± 5
Elite athletes (%)	54	33	50	0	33	0
Controlled trials (%)	46	11	85	100	100	100
Blind trials (%)	0	0	0	20	67	14
Males (%)	84	61	80	72	81	90
Competitive phase (%)	31	33	10	20	67	14
Phase unknown (%)	54	56	90	60	0	43
Maximal tests (%)	78	100	85	100	71	92
Altitude level (m)	2030 ± 410	2400 ± 290	2890 ± 420	4530 ± 840	6000	2750 ± 310
Hours of hypoxia per day	24	~18–24	11 ± 3			
Minutes of hypoxia per day				210 ± 84	40 ± 9	47 ± 48
Days of exposure	23 ± 6	27 ± 1	18 ± 7	9 ± 3	16 ± 2	14 ± 4

Continued next page

Table II. Contd

Effect	Natural altitude protocols		Artificial altitude protocols			
	live-high train-high	live-high train-low	live-high 8–18 h/d, continuous, train-low	live-high 1.5–5 h/d, continuous, train-low	live-high <1.5 h/d, intermittent, train-low	live-low train-high 0.5–2 h/d
Total period of treatment (d)	23 ± 6	27 ± 1	19 ± 7	14 ± 5	20 ± 3	30 ± 13
Exposure/treatment ratio (%)	100	100	96 ± 7	76 ± 32	82 ± 8	55 ± 27
Training intensity (1–4)						2.3 ± 1.1
Post-exposure test day[d]	9.1 ×/÷ 1.9	5.4 ×/÷ 2.2	2.2 ×/÷ 2.7	4.4 ×/÷ 1.9	6.4 ×/÷ 2.3	2.8 ×/÷ 1.9
Duration of maximal tests[d] (min)	6.9 ×/÷ 2.4	*11 ×/÷ 1.3*	5.6 ×/÷ 2.3	5.2 ×/÷ 3.0	6.1 ×/÷ 2.5	3.9 ×/÷ 4.9
Effects of study characteristics (%); ±90% CL						
Subelite-elite	(0.7; ±3.8)	(0.3; ±2.2)	(0.8; ±3.2)		**2.4; ±2.8**	
Uncontrolled-controlled	3.3; ±3.6	*–2.6; ±3.0*	(–1.6; ±3.4)			
Blind-not blind					(–1.4; ±4.5)	
Female-male		(–0.3; ±3.9)				
Competitive-unknown phase	(0.5; ±3.8)					
Submaximal-maximal test	(0.0; ±1.6)		*–3.3; ±2.4*		(–0.3; ±1.8)	(1.4; ±3.2)
1 SD altitude level	**1.2; ±1.6**	–0.1; ±1.0	**1.5; ±2.5**	**–2.3; ±2.5**		(–0.9 ±2.5)
1 SD hours hypoxia			0.8; ±1.8			
1 SD minutes hypoxia					(0.4; ±2.3)	
1 SD days exposure	*–1.8; ±1.7*		*–1.0; ±1.7*			**2.4; ±2.5**
1 SD exposure/treatment ratio					0.6; ±1.2	
1 SD training intensity						(–1.2; ±2.5)
1 SD post-exposure test day	0.5; ±0.7	–0.2; ±0.3	(0.1; ±2.1)	–0.5; ±0.8	–0.4; ±0.6	**1.2; ±1.5**
1 SD duration of max. test	**3.0; ±2.5**		–0.9; ±1.2	–0.3; ±0.6	0.6; ±1.3	–0.2; ±1.1
Random variation (%); ±90% CL or ×/÷90% CL factor						
Between-study SD	2.7; ±2.3	1.3; ±1.3	1.0; ±1.9	2.2; ±3.5	–0.6; ±0.9	2.4; ±3.1
Standard error of measurement	2.4; ×/÷1.7	0.7; ×/÷2.2	2.2; ×/÷1.9	1.2; ×/÷ 1.9	3.2; ×/÷1.3	2.8; ×/÷1.5

a Effects are the means predicted for controlled trials and maximal tests, but are otherwise evaluated at the mean values of the study characteristics for which effects are shown.

b 90% CL: subtract and add this number to the observed effect to obtain the 90% CL for the true (large-sample) effect.

c Effects are the predicted means in maximal tests adjusted to ±1 SD away from the mean for selected study characteristics shown.

d SD shown as ×/÷ factor derived from log-transformed times.

CL = confidence limits.

In most models, it was possible to include a random effect to estimate pure between-study variation in the effect of the treatment, expressed as a standard deviation. In principle, this measure of between-study variation is free of sampling variation arising from error of measurement in the dependent variable, but use of sample size as the weighting factor does not produce clean partitioning of random error into pure between-study variation and residual error. The standard deviation representing the residual error in such models is the standard error of a study-estimate with the mean sample size (n) of the meta-analysed estimates; this standard error was multiplied by $\sqrt{(n/8)}$ to provide an estimate of the mean standard error of measurement of the dependent variable. When there were insufficient study-estimates to include a pure between-study random effect, the residual random effect is shown as the between-study standard deviation.

For each outcome measure, a novel funnel plot of the inverse of an estimate's weighting factor (y-axis) versus the value of the estimate's random effect (x-axis) was examined qualitatively for evidence of outliers (points judged visually to be more than about 4 SDs of horizontal scatter away from the centre of the plot) and publication bias towards positive effects (positive trend in the scatter). This procedure did not result in exclusion of any estimates.

We reported uncertainty in the meta-analysed estimates as 90% confidence limits, and we made probabilistic magnitude-based inferences about the true (large-sample) values of outcomes, as described elsewhere.[76] In brief, an outcome was deemed unclear if its confidence interval overlapped the thresholds for smallest worthwhile positive and negative effects; equivalently, effects were unclear if chances of the true value being substantially positive and negative were both >5%. The magnitude of a clear effect was reported as the magnitude of its observed value, sometimes with an estimate of the probability the true value was substantial. The probabilities for each meta-analysed effect and for pairwise comparisons of effects were derived using a published spreadsheet.[77] The thresholds for smallest effects on performance were assumed to be ±1%, which is an approximate average across a range of sports.[78-80] Smallest effects on $\dot{V}O_{2max}$ and exercise economy were also assumed to be ±1% because the relationship between these measures and endurance performance[6] implies that, other things being equal, a 1% change in either of these measures would result in a similar change in performance. We also assumed a smallest effect of 1% for haemoglobin or red-cell mass because $\dot{V}O_{2max}$ is effectively proportional to haemoglobin mass in a cross-sectional study of athletes.[81] For haemoglobin and peak lactate concentration, there is no direct relationship with performance; effects were therefore standardized by dividing by the mean between-subject standard deviation of these variables in the studies that contributed to their meta-analyses, and a modified Cohen scale was used to make inferences.[82]

2. Results

2.1 Exercise Performance

The meta-analysed outcomes for the six protocols of natural and artificial altitude are shown in table II. Substantial enhancement of power output in subelite athletes was very likely with artificial brief intermittent LHTL, likely with LHTL, possible with artificial long continuous LHTL, but unclear for LHTH, artificial brief continuous LHTL and LLTH. Comparisons between the protocols for subelite athletes revealed that LHTL was likely better than all protocols, with the exception of artificial brief intermittent LHTL, where the difference was unclear. Artificial brief intermittent LHTL was possibly better than artificial long continuous LHTL, artificial brief continuous LHTL and LLTH. All other differences between protocols were unclear. Enhancements of mean power in elite athletes were likely with LHTL, but unclear for all other protocols. In comparison with the other protocols in elite athletes, LHTL was likely better than artificial long continuous LHTL and artificial brief intermittent LHTL. All other differences between protocols were unclear.

Several of the study characteristics listed in table II moderated the effects of hypoxia; performance was better in controlled relative to uncontrolled studies for LHTL, but the opposite was observed for LHTH; subelite athletes had a clear enhancement in performance relative to elite athletes with artificial brief intermittent LHTL, and submaximal exercise performance was clearly impaired relative to maximal with artificial long continuous LHTL. Effects for blinding, competitive phase and sex were unclear for the few protocols where these effects could be estimated. Post-exposure test day had a substantial clear positive linear effect for LLTH and trivial or unclear effects with the other protocols. Quadratic or cubic effects of post-exposure test day (not shown in table II) could not be modelled with the two shortest protocols of artificial altitude, and the polynomials revealed little curvature with LHTL (<0.3% over $\times/\div SD^2$ either side of the mean time). However, relative to the effect at the mean post-test time, LHTH showed some evidence of enhancement at very short times (1.8% at $\div SD^2$ or ~2.5 days; 90% confidence limits ±4.7%) followed by impairment (−1.5% at ÷SD or 5 days; ±1.9%), enhancement (1.4% at ×SD or 17 days; ±1.9%) and impairment (−2.3% at $\times SD^2$ or 33 days; ±8.9%); artificial long continuous LHTL showed a peak at the mean post-test time with a relative impairment of 1–2% (~±4.5%) either side of the mean (at ×/÷SD or 1 and 6 days); and LLTH showed a trough at the mean time with relative enhancements at ÷SD or 1.5 days (3.8%; ±5.3%) and at ×SD or 5 days (1.0%; ±1.4%).

The moderating effect of study characteristics provides an avenue for enhancing each protocol, as shown in table II for the effects on performance after changing selected characteristics by ± or ×/÷1 SD. Improvements in power output were observed in subelite athletes for all protocols after these theoretical enhancements, the increase ranging from 0.4% for LHTL to 5.9% for LLTH. The resulting effects were all clearly beneficial for subelite athletes, but beneficial effects for elite athletes were clear only for LHTH and LHTL. Alterations to the altitude level, days of exposure and daily exposure hours had the biggest contribution to the enhanced protocols, whereas effects for other characteristics were generally trivial or unclear. Modifying test duration by one SD would also have produced substantial enhancements in performance, especially for LHTH, but this characteristic was not included because the mean duration of tests was reasonably similar across the protocols, and changing the performance test does not represent a change to an exposure protocol.

2.2 Physiological Measures

The meta-analysed effects on sea-level $\dot{V}O_{2max}$ are shown in table III. There was a very likely enhancement with LHTH and a possible enhancement with LLTH in subelite athletes. The trivial effect for artificial LHTL with predominantly subelite athletes is very unlikely to have arisen from a substantial true positive effect. The unclear effects for the remaining two artificial protocols represent changes in $\dot{V}O_{2max}$ that were either unlikely to be positive (brief continuous LHTL) for subelite athletes or were possibly positive (brief intermittent LHTL) for predominantly subelite athletes. For elite athletes, there was a possible 'impairment' with LHTH, but an unclear effect for LHTL. It was not possible to estimate effects for elite athletes alone in the other protocols.

Study characteristics moderating $\dot{V}O_{2max}$ are also shown in table III. The most interesting effect of characteristics with the natural protocols was the increase in $\dot{V}O_{2max}$ with increasing time post-exposure (clear for LHTH, unclear for LHTL), indicating that there is more benefit at least for $\dot{V}O_{2max}$ around 2 weeks after the intervention period. The trivial effect in artificial long continuous LHTL can be converted into a positive effect by increasing the hours of exposure; there is also a possibility of less benefit from 'more' days of exposure, even though the mean number of days of exposure is already about a week less than for the natural protocols. A reduction in training intensity with LLTH would promote a further increase in $\dot{V}O_{2max}$. The remaining effects of study characteristics on $\dot{V}O_{2max}$ were unclear.

Table III. Meta-analysis of effects on sea-level maximal oxygen uptake following adaptation to hypoxia experienced in studies with various protocols of natural and artificial altitude. Effects of mean protocol are those predicted for controlled trials. Effects in parentheses are unclear (>5% chance of increase and >5% chance of decrease); otherwise **bold** indicates ≥50% chance of increase, *italic* indicates ≥50% chance of decrease, and plain font indicates ≥50% chance of a trivial effect. These probabilistic outcomes are computed with reference to a smallest important change of 1%

Effect	Natural altitude		Artificial altitude			
	live-high train-high	live-high train-low	continuous long hypoxia (8–18 h/d), train-low	continuous brief hypoxia (1.5–5 h/d), train-low	intermittent brief hypoxia (<1.5 h/d), train-low	live-low, train-high (0.5–2 h/d)
Effect of mean protocol[a] (%); ±90% CL[b]						
Elite	*−1.5; ±2.0*	(6.4; ±11.2)	−0.5; ±1.4		(0.1; ±2.8)	
Subelite	**4.3; ±2.6**	(6.4; ±9.4)		(−1.1; ±3.5)		**1.1; ±2.0**[e]
Study characteristics (mean ± SD)[c]						
References	12	5	5	4	3	8
Study groups	15	9	6	5	3	8
Study estimates	20	10	7	6	5	10
Subjects per estimate	15 ± 7	12 ± 6	20 ± 10	15 ± 5	19 ± 5	16 ± 5
Effective subjects per estimate[d]	33 ± 19	41 ± 11	20 ± 10	15 ± 5	19 ± 5	16 ± 5
Elite athletes (%)	57	33	33	0	33	0
Controlled trials (%)	43	11	100	100	100	100
Blind trials (%)	0	0	0	20	33	13
Males (%)	87	61	75	72	100	91
Competitive phase (%)	29	33	0	20	100	13
Phase unknown (%)	57	56	100	60	0	50
Altitude level (m)	1990 ± 400	2400 ± 290	2680 ± 160	4530 ± 880	6000	2970 ± 680
Hours of hypoxia per day			10 ± 2			
Minutes of hypoxia per day				210 ± 90	35 ± 8	45 ± 46
Days of exposure	23 ± 6	27 ± 1	18 ± 6	9 ± 3	17 ± 3	14 ± 4
Total period of treatment (d)	23 ± 6	27 ± 1	19 ± 7	14 ± 5	21 ± 4	28 ± 14
Exposure/treatment ratio (%)	100	100	96 ± 9	76 ± 33	78 ± 1	59 ± 27
Training intensity (1–4)						2.1 ± 1.1
Post-exposure test day[c]	8.0 ×/÷ 1.8	4.8 ×/÷ 2.1	1.2 ×/÷ 2.3	4.4 ×/÷ 2.0	4.7 ×/÷ 2.5	2.9 ×/÷ 2.2
Effects of study characteristics (%); ±90% CL						
Uncontrolled-controlled	0.3; ±2.4	(−2.7; ±9.3)				
Competitive-unknown phase	(1.3; ±2.7)					
Subelite-elite	**5.5; ±2.4**	(−0.0; ±5.2)				
1 SD altitude level	0.3; ±1.2					(0.0; ±2.3)

Continued next page

Table III. Contd

Effect	Natural altitude		Artificial altitude			
	live-high train-high	live-high train-low	continuous long hypoxia (8–18 h/d), train-low	continuous brief hypoxia (1.5–5 h/d), train-low	intermittent brief hypoxia (<1.5 h/d), train-low	live-low, train-high (0.5–2 h/d)
1 SD hours exposure			**2.5; ±1.9**			
1 SD days exposure	0.5; ±1.3		−0.9; ±1.7	(1.1; ±4.3)		(1.0; ±2.2)
1 SD training intensity						*−1.2; ±2.1*
1 SD post-exposure test day	**1.0; ±1.0**	(1.1; ±2.8)			(−0.5; ±3.3)	
Random variation (%); ±90% CL or ×/÷90% CL factor						
Between-study SD	1.8; ±2.4	3.8; ×/÷1.7[e]	1.7; ×/÷1.9[e]	3.3; ×/÷2.5[e]	2.6; ×/÷2.2[e]	2.1; ±2.8
Standard error of measurement	2.9; ×/÷1.8					2.5; ×/÷2.6

a Effects are the means predicted for controlled trials, but otherwise evaluated at the mean values of the study characteristics for which effects are shown.

b 90% CL: subtract and add this number to the observed effect to obtain the 90% CL for the true (large-sample) effect.

c SD shown as ×/÷ factor derived from log-transformed times.

d Derived by adjusting all sample sizes to those of controlled trials with equal numbers in control and experimental groups.

e Insufficient within-study clusters to estimate error of measurement; between-study SD includes within-study sampling variation.

CL = confidence limits.

Haemoglobin mass (including red-cell mass) and exercise economy were meta-analysed for all studies collectively because of the lack of study estimates. Effects for haemoglobin mass were unclear, but an increase in exposure days and possibly an increase in altitude would produce a clear increase, whereas delaying the test day by >1 SD (>10 days) would offset the increase. The effect on exercise economy was trivial, but a substantial increase could accrue from reducing exposure days and increasing altitude (table IV).

Haemoglobin concentration and peak lactate could be meta-analysed only for LHTH and brief intermittent artificial LHTL. For the interpretation of magnitude, the average pre-test between-subject standard deviation for haemoglobin concentration was 6.2%, while that for peak lactate was 21%. Haemoglobin concentration demonstrated a likely moderate increase for LHTH and a possible small increase for artificial brief intermittent LHTL. The moderating effect of post-exercise test day shows that the increase in haemoglobin concentration was lost 3–4 weeks after exposure. The effect for peak lactate was unclear with LHTH, but an increase in altitude would produce a clear small to moderate decrease, whereas delaying the test day would produce a similar (but unexpected) decrease. Peak lactate showed a trivial decrease for artificial LHTL. The effect for peak lactate in artificial brief intermittent LHTL was trivial, but the uncertainty allows for the possibility of a small negative true effect.

Effects for other physiological measures that could not be meta-analysed due to insufficient data are shown in figure 1. Erythropoietin was elevated during the hypoxic interventions and possibly showed a small elevation afterwards. Reticulocytes appeared to be elevated in a few studies during the intervention. The scatter in the plot for ferritin makes any conclusion about trend difficult.

Plots of performance versus $\dot{V}O_{2max}$, haemoglobin or red-cell mass and exercise economy are shown in figure 2. An estimate of the strength of the relationship between performance and each of these variables (in units of percentage change

Table IV. Meta-analysis of effects on sea-level haemoglobin (Hb) or red-cell mass (Hb mass), exercise economy, Hb concentration, and peak lactate in an exercise test following adaptation to hypoxia experienced in studies with various protocols of natural and artificial altitude. Effects in parentheses are unclear (>5% chance of increase and >5% chance of decrease); otherwise **bold** indicates ≥50% chance of increase, *italic* indicates ≥50% chance of decrease, and plain font indicates ≥50% chance of a trivial effect. These probabilistic outcomes are computed with reference to a smallest important change of 1% for Hb mass, 1% for economy, and 0.20 of baseline between-subject SD for Hb concentration and peak lactate

Effect	Hb mass, where measured[a]	Economy, where measured[b]	Hb concentration		Peak lactate	
			LHTH	intermittent brief hypoxia, train low	LHTH	intermittent brief hypoxia, train low
Effect of mean protocol[c](%); ±90%CL[d]	(1.3; ±2.4)	0.4; ±1.3	**4.8; ±2.7**	**2.3; ±1.2**	(0.7; ±5.7)	–3.5; ±4.7
Study characteristics (mean ± SD)						
References	12	14	5	4	5	5
Study groups	14	15	7	4	7	7
Study estimates	18	19	8	8	9	14
Subjects/estimate	15 ± 7	19 ± 5	16 ± 9	22 ± 5	19 ± 8	24 ± 3
Effective subjects/estimate	25 ± 9	31 ± 27	32 ± 11	22 ± 5	35 ± 23	24 ± 3
Elite athletes (%)	46	33	57	20	43	0
Controlled trials (%)	62	80	43	100	57	100
Blind trials (%)	0	7	0	60	0	50
Males (%)	74	91	83	92	82	90
Competitive phase (%)	15	53	29	80	29	75
Phase unknown (%)	62	40	43	0	43	0
Altitude level (m)	2540 ± 970	3410 ± 1460	1900 ± 280	6000	1990 ± 320	6000
Minutes of hypoxia per day				37 ± 7		35 ± 6
Days of exposure	21 ± 7	20 ± 6	24 ± 5	16 ± 2	22 ± 6	15
Total period of treatment (d)	21 ± 7	24 ± 6	24 ± 5	20 ± 4	22 ± 6	18 ± 2
Exposure/treatment ratio (%)	100	86 ± 22	100	83 ± 9	100	84 ± 9
Post-exposure test day[e]	3.9 ×/÷ 2.6	3.3 ×/÷ 2.8	9.1 ×/÷ 2.1	4.3 ×/÷ 2.3	8.3 ×/÷ 2.2	5.9 ×/÷ 2.3
Effects of study characteristics (%); ±90% CL						
1 SD altitude level	(1.5; ±2.6)	0.6; ±1.6	(–1.8; ±4.0[e])		*–12.4; ±7.0*	
1 SD exposure days	**2.7; ±2.7**	–0.8 ±1.6			(1.4; ±7.7)	
1 SD post-exposure test day	–0.9; ±1.0	(0.1; ±1.4)	*–3.3; ±3.9*	0.6; ±1.3	*–10.5; ±7.7*	–0.6; ±2.1
Random variation (%); ±90% CL or ×/÷90% CL factor						
Between-study SD	4.6; ±2.2	–1.0; ±2.5	3.6; ×/÷1.8[f]	1.7; ×/÷1.7[f]	8.7; ×/÷1.7[f]	3.6; ±4.8
Standard error of measurement	2.1; ×/÷1.9	6.0; ×/÷1.7				7.3; ×/÷1.5

a Number of estimates: LHTH, 10; LHTL, 3; artificial long continuous LHTL, 3; artificial brief continuous LHTL, 2.

b Number of estimates: LHTH, 4; LHTL, 3; artificial long continuous LHTL, 3; artificial brief continuous LHTL, 3 artificial brief intermittent LHTL, 5; LLTH, 1.

c Effects are the predicted means evaluated at the mean values of the study characteristics for which effects are shown.

d 90% CL: subtract and add this number to the observed effect to obtain the 90% CL for the true (large-sample) effect.

e SD shown as ×/÷ factor derived from log-transformed times.

f Insufficient within-study clusters to estimate error of measurement; between-study SD includes within-study sampling variation.

CL = confidence limits; **LHTH** = live-high train-high; **LHTL** = live-high train-low; **LLTH** = live-low train-high.

in performance per percentage change in the variable) is provided by the slope of the regression line for each protocol (not shown in the figure). The only clear slopes were for $\dot{V}O_{2max}$ with LHTH (0.49 %/%; 90% confidence limits ±0.29%/%) and LHTL (0.22; ±0.13%/%).

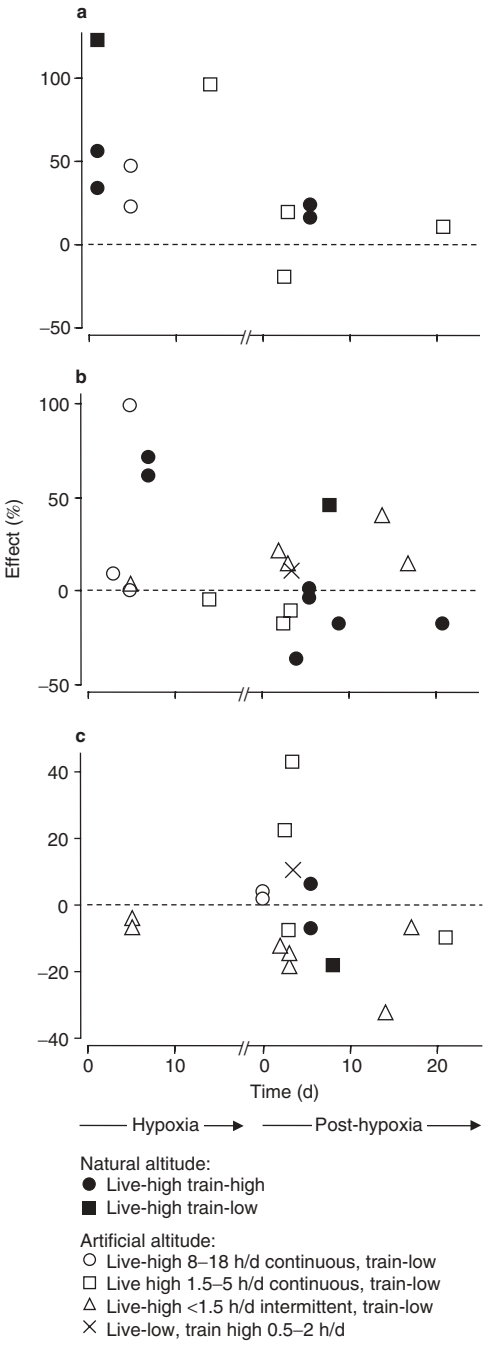

Fig. 1. Individual study-estimates of effects on (**a**) erythropoietin, (**b**) reticulocytes and (**c**) ferritin sampled in blood during and following exposure to hypoxia with the various protocols of natural and artificial altitude.

3. Discussion

In this first meta-analysis of sea-level exercise performance following adaptation to hypoxic exposure, we observed clear enhancements in endurance power output of 1–4% in subelite athletes with LHTL and with two of the artificial-altitude protocols (long continuous and brief intermittent LHTL). In elite athletes, the enhancements were clear only with LHTL. Modification of study characteristics might result in clear enhancements of 3–7% with all protocols in subelite athletes, but effects in elite athletes would be clear only for LHTH and LHTL.

Following the development of the LHTL approach, the use of LHTH has received little support from sport scientists. There is enough uncertainty in our estimates of the effect of LHTH to allow for enhancements in elite and subelite athletes with this protocol. Furthermore, our estimates are for controlled trials, whereas athletes in an altitude camp would experience the equivalent of an uncontrolled trial, giving a possible further increase of ~3% (table II). The LHTH protocol also showed effects of post-exposure test day consistent with anecdotal reports of coaches that performance is enhanced immediately after altitude and peaks again several weeks later. Taken together, these results provide reasonable support for what is still a widely accepted practice among many elite coaches and athletes. LHTH was also one of only two protocols that produced clear enhancements in endurance performance for elite athletes with appropriate manipulation of study characteristics. These moderating effects show that it may be better for athletes to go to higher altitudes (~2400 m) for shorter periods (~16 days) around 2–3 weeks before an important competition.

Our results provide good evidence for the effectiveness of LHTL, which was clearly better than all but one protocol in subelite athletes (brief intermittent LHTL) and elite athletes (LHTH). The only moderating effect of study characteristics with LHTL was unexpected: uncontrolled trials showed a clear negative effect relative to controlled trials. According to conventional wisdom, uncontrolled studies should show 'larger'

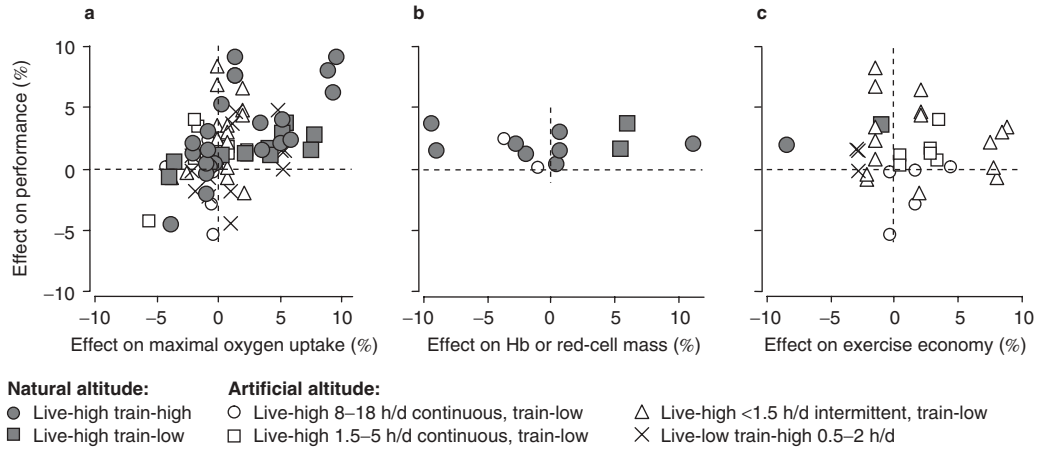

Fig. 2. Individual study-estimates of effects on performance plotted against maximal oxygen uptake, haemoglobin (Hb) or red-cell mass, and exercise economy with the various protocols of natural and artificial altitude.

enhancements due to the so-called 'training-camp effect', which in principle is adjusted for in a controlled trial. What may happen in reality is that subjects in the control group of a controlled trial experience less of the training-camp effect, because they do not train as hard. There could also be a contribution from a 'nocebo effect', whereby subjects in a control group perform worse, because they know they are in the control group. There is evidence of a nocebo effect in the classic natural LHTL study of Levine and Stray-Gundersen[2] that is especially clear when the data are presented graphically as percentage changes (see figure 1 in Baker and Hopkins[83]). Indeed, data for the effect of uncontrolled versus controlled LHTL studies came entirely from this study. Therefore, our meta-analysed effect of ~4% for controlled studies needs to be interpreted with caution. When performance is predicted for uncontrolled studies (as previously mentioned, the way athletes train), the effect becomes a more realistic ~1.5%. The only design that avoids the nocebo problem is a blind trial, which is not possible with natural LHTL. Further research with controlled trials is warranted to assess the potential of LHTL.

Artificial LHTL with long continuous exposures was developed to simulate LHTL, and our analysis provides some support for its efficacy. The limitation with this approach appears to be insufficient exposure to hypoxia because the moderating effects of study characteristics show that the effect on performance can be increased by increasing altitude and adding daily exposure hours. This result is consistent with the suggestions of researchers who believe that at least 12 hours of daily exposure is critical for the success of this protocol.[7,84] Another substantial moderating effect was a reduction in performance with increasing days of exposure, similar to that with LHTH. This result for both protocols seems counter-intuitive, although a ready explanation is a short-acting placebo effect. The only other moderating effect was a substantial downward adjustment for submaximal performance, which again implicates a placebo effect. More studies are needed to clarify the role of placebo effects with this and other protocols.

At the opposite end of the spectrum of daily hypoxic exposure, artificial LHTL with brief intermittent exposures was one of the best protocols in subelite athletes. The moderating effects of study characteristics provided only marginal improvements of 1%, mainly through maximizing the exposure days in the intervention period. The equivalent altitude of this protocol is already at the limit for ethical approval, so there is no option to investigate higher altitudes. Alteration

of the hypoxic and normoxic intervals is a possible avenue for improvement, although we have found no clear difference between the effects of 3- and 5-minute intervals.[60] The clear difference between the effect on subelite and elite athletes suggests that the waves of hypoxia are less effective in elite athletes, possibly because elite athletes experience more hypoxia in their muscles from higher intensities of training compared with subelite athletes.

With the remaining two forms of artificial altitude exposure, the uncertainty in the meta-analysed estimates was too large for their trivial magnitudes to be clear, although clear enhancements were possible with adjustment of appropriate study characteristics. With brief continuous LHTL, the average altitude appears to have been too high, since a reduction in altitude by 1 SD could produce a substantial enhancement in performance. Reducing the altitude may seem an implausible way to enhance this protocol, but the reduction by 1 SD would bring the altitude to ~3700 m, which is still well above that of the other continuous protocols and which could conceivably provide a sufficient hypoxic stimulus without the negative sequelae of continuous exposure to high altitude. A reduction in altitude along with a reduction in training intensity would also enhance performance with LLTH, but the main enhancement for this protocol would come from the more reasonable strategy of increasing days of exposure. LLTH also showed evidence that performance could be better either side of the mean post-exposure test day (~3 days), but it seems to us that this protocol is the least likely to produce performance enhancement.

A study characteristic not included in the above discussion of the individual protocols was test duration, because altering this characteristic does not alter the exposure protocol. There are nevertheless implications for the effects on aerobic versus anaerobic performance. Our results demonstrate that performance could be improved in LHTH and brief intermittent LHTL with tests of longer duration. In all other protocols, performance could be better by a trivial margin for shorter tests. The average test duration in all protocols was 4–11 minutes, making all tests highly aerobic, but with only one of the protocols (LLTH) would a 1-SD reduction in test duration make the tests substantially anaerobic. More studies with shorter tests are needed to clarify the effect on anaerobic performance.

Insights into the practical application of the findings of the meta-analyses can also be gleaned from a consideration of the between-study standard deviations (table II). These standard deviations represent unexplained variation in the mean effect of the protocol from study to study; as such, their magnitude is the typical deviation from the meta-analysed mean effects that a researcher or practitioner can expect to experience in another study using the mean protocol with a group or squad. For natural and artificial brief intermittent LHTL, these standard deviations in combination with the uncertainties in the mean effects imply that most researchers and practitioners will observe substantial enhancements in performance with a group or squad of subelite athletes. A beneficial outcome is less certain for elite athletes with the natural protocols and for subelite athletes with artificial long continuous LHTL; for the remaining protocols with subelite or elite athletes the outcomes could be good, bad or indifferent. However, if the enhanced protocols are as good as shown, the influence of the between-study standard deviation could be nullified for all protocols.

The standard errors of measurement estimated from the meta-analyses (table II) do not have an immediate practical application, but they do provide evidence that the uncertainties in the meta-analysed mean effects and in the moderating effects of study characteristics are trustworthy. These uncertainties are estimated from a combination of the between-study standard deviations and the standard errors of measurement, so it is important that the standard errors of measurement estimated from the meta-analysis are realistic. The low value for natural LHTL (0.7%) is a reflection of the fact that almost all of the performance tests in these studies were time trials with runners. This value and the other values for error of measurement, given their uncertainties, are within the normal range for tests of endurance performance.[73]

It is important to understand that some individual athletes may obtain no benefit or even impairment in performance from adaptation to hypoxia, even with those protocols that are clearly beneficial. Meta-analysis cannot address the question of individual responses to treatments until researchers provide complete inferential information about experimental and control groups in the form of confidence limits, exact p-values, or best of all, standard deviations of change scores. Such information would also allow the use of the inverse of sampling variance instead of sample size as a weighting factor in the meta-analysis, which would result in more trustworthy and probably narrower uncertainties in the meta-analysed mean and between-study standard deviations.

Turning to the analysis of potential mechanism variables, it is clear from the findings in table III that adaptation to hypoxia can result in enhancements in maximal oxygen uptake. The usual mechanism suggested for an increase in this variable is erythropoiesis, which would effect a change in haemoglobin or red cell mass with a resulting increase in blood volume, cardiac output or oxygen-carrying capacity. Our meta-analyses provide limited evidence for this mechanism: the meta-analysed effect on haemoglobin mass was unclear on average, although extra exposure to hypoxia and a higher altitude level could result in a substantial increase. The meta-analysed effects on haemoglobin concentration provide some additional indirect evidence for an increase in haemoglobin mass, but an alternative explanation for the increase in haemoglobin concentration often mentioned by researchers is a dehydrating effect of acclimatization to altitude.[1] Direct evidence of erythropoiesis from levels of erythropoietin and reticulocytes could not be provided by meta-analysis, due to insufficient data, but it is reasonably clear from figure 1 that these variables increase transiently to some extent in some studies. Any erythropoiesis that did occur was not accompanied by clear reductions in ferritin, although supplementation with iron in most studies would probably offset any reduction.

Do the changes in $\dot{V}O_{2max}$ mediate the changes in performance? The pattern of the effects on $\dot{V}O_{2max}$ in table III across different protocols for elite and subelite athletes and for the moderating effects of study characteristics does not mirror closely the effects on performance in table II. On the other hand, the relationships (slopes) between changes in $\dot{V}O_{2max}$ and performance were clear for the natural-altitude protocols and were of a magnitude that might be expected if $\dot{V}O_{2max}$ was a primary mediator, given the attenuating effects that error of measurement in this variable would have on the slopes. Unfortunately, blinding was not possible with these protocols, and therefore placebo and nocebo effects may have contributed to the relationships. A positive relationship between changes in haemoglobin mass and changes in performance – the expected outcome if the changes in $\dot{V}O_{2max}$ were mediated by erythropoiesis was not observed (figure 2), although error of measurement with haemoglobin mass (which manifests as a large between-study coefficient of variation, table IV) could also attenuate a true substantial relationship. Thus, our analyses have not resolved the issue of whether $\dot{V}O_{2max}$ is the primary mediator of performance following adaptation to hypoxia.

There were insufficient data to meta-analyse the effects of exercise economy for each protocol, but a single analysis for all protocols and the relationship between exercise economy and sea-level performance (figure 2) provided little evidence for this mechanism. The only other physiological variable we meta-analysed, peak lactate concentration, is not a contender as a primary mechanism of performance enhancement, but an increase in peak blood lactate would indirectly implicate buffering capacity. However, placebo and nocebo effects on performance in altitude and control groups could also lead to an increase in peak lactate. For the two protocols we meta-analysed, a substantial increase in peak blood lactate was either unlikely (LHTH) or very unlikely (brief intermittent LHTL), so an increase in buffering capacity is presumably not involved with adaptation to these protocols. Other mechanisms therefore need to be identified, particularly for the artificial LHTL protocols, where gains in performance appear to be due at least partly to placebo or nocebo effects and where an increase in $\dot{V}O_{2max}$ may not contribute.

The suggestion of a change in cardiovascular regulation resulting in more cardiac output to exercising muscle[4] is plausible, but will be hard to investigate.

4. Conclusions

Meta-analysis cannot adjust for the confounding effects of unknown or unquantified study characteristics. Furthermore, the simplistic nature of linear modelling, the exclusion of interactions between predictors, and the inevitable presence of substantial random error and systematic bias with some predictors all conspire to prevent the meta-analytic model from fully accounting for confounding effects even of the study characteristics included in the model. Nevertheless, our method of estimation of confidence intervals based on weighting by sample size is conservative, so our analyses must provide some evidence of the efficacy of adaptation to hypoxia for physical performance. Subelite athletes can experience endurance performance enhancements with adaptation to natural altitude exposure and to brief intermittent and long continuous protocols of artificial altitude exposure. For elite athletes, enhancements in endurance performance were possible only with the natural LHTL protocol. The enhancements with natural altitude could be mediated in part by $\dot{V}O_{2max}$, but placebo effects, nocebo effects, training-camp effects and other mechanisms may be involved with these and the artificial protocols. Perhaps the most important outcomes of our analyses are the suggestions for enhancement of the protocols, some of which should be the focus of future research using double-blind designs, performance measures with smaller errors of measurement, and putative physiological mediators. Reviewers and editors should ensure that studies accepted for publication contain complete inferential information about the effects in treatment and control groups.

Acknowledgements

The literature reviews in the theses of Erica Hinckson and Matt Wood provided a valuable starting point for this review. Chris Gore provided useful publication lists and feedback on a draft version. The only funding for this study was provided by our institutional employer as salaries. There are no conflicts of interest.

References

1. Wilber RL. Current trends in altitude training. Sports Med 2001; 31: 249-65
2. Levine BD, Stray-Gundersen J. 'Living high-training low': effect of moderate-altitude acclimatization with low-altitude training on performance. J Appl Physiol 1997; 83: 102-12
3. Gore CJ, Clark SA, Saunders PU. Nonhematological mechanisms of improved sea-level performance after hypoxic exposure. Med Sci Sports Exerc 2007; 39: 1600-9
4. Gore CJ, Hopkins WG. Counterpoint: positive effects of intermittent hypoxia (live high-train low) on exercise performance are not mediated primarily by augmented red cell volume. J Appl Physiol 2005; 99: 2055-7
5. Levine BD, Stray-Gundersen J. Point: positive effects of intermittent hypoxia (live high: train low) on exercise performance are mediated primarily by augmented red cell volume. J Appl Physiol 2005; 99: 2053-5
6. di Prampero PE. The energy cost of human locomotion on land and in water. Int J Sports Med 1986; 7: 55-72
7. Levine BD. Intermittent hypoxic training: fact and fancy. High Alt Med Biol 2002; 3: 177-93
8. Stray-Gundersen J, Chapman RF, Levine BD. 'Living high-training low' altitude training improves sea level performance in male and female elite runners. J Appl Physiol 2001; 91: 1113-20
9. Wood MR, Dowson MN, Hopkins WG. Running performance after adaptation to acutely intermittent hypoxia. Eur J Sport Sci 2006; 6: 163-72
10. Brugniaux JV, Schmitt L, Robach P, et al. Eighteen days of 'living high, training low' stimulate erythropoiesis and enhance aerobic performance in elite middle-distance runners. J Appl Physiol 2006; 100: 203-11
11. Robach P, Schmitt L, Brugniaux JV, et al. Living high-training low: effect on erythropoiesis and maximal aerobic performance in elite Nordic skiers. Eur J Appl Physiol 2006; 97: 695-705
12. Robach P, Schmitt L, Brugniaux JV, et al. Living high-training low: effect on erythropoiesis and aerobic performance in highly-trained swimmers. Eur J Appl Physiol 2006; 96: 423-33
13. Schmitt L, Millet G, Robach P, et al. Influence of 'living high-training low' on aerobic performance and economy of work in elite athletes. Eur J Appl Physiol 2006; 97: 627-36
14. Adams WC, Bernauer EM, Dill DB, et al. Effects of equivalent sea level and altitude training on $\dot{V}O_{2max}$ and running performance. J Appl Physiol 1975; 39: 262-6
15. Balke B, Nagle FJ, Daniels J. Altitude and maximum performance in work and sports activity. JAMA 1965; 194: 646-9
16. Bushkirk ER, Kollias J, Akers RF, et al. Maximal performance at altitude and on return from altitude in conditioned runners. J Appl Physiol 1967; 23: 259-66

17. Casas M, Casas H, Pages T, et al. Intermittent hypobaric hypoxia induces altitude acclimation and improves the lactate threshold. Aviat Space Environ Med 2000; 71: 125-30
18. Chapman RF, Stray-Gundersen J, Levine BD. Individual variation in response to altitude training. J Appl Physiol 1998; 85: 1448-56
19. Daniels J, Oldridge N. The effects of alternate exposure of altitude and sea level on world-class middle-distance runners. Med Sci Sports Exerc 1970; 2: 107-12
20. Hellemans J. Intermittent hypoxic training: a pilot study. Proceedings of the Second Annual International Altitude Training Symposium. Flagstaff: 1999: 145-54
21. Klausen K, Robinson S, Micahel ED, et al. Effect of high altitude on maximal working capacity. J Appl Physiol 1966; 21: 1191-4
22. Piehl Aulin K, Svedenhag J, Wide L, et al. Short-term intermittent normobaric hypoxia-hematological, physiological and mental effects. Scand J Med Sci Sports 1998; 8: 132-7
23. Rodriguez FA, Casas H, Casas M, et al. Intermittent hypobaric hypoxia stimulates erythropoiesis and improves aerobic capacity. Med Sci Sports Exerc 1999; 31: 264-8
24. Rodriguez FA, Murio J, Ventura JL. Effects of intermittent hypobaric hypoxia and altitude training on physiological and performance parameters in swimmers [abstract]. Med Sci Sports Exerc 2003; 35: S115
25. Telford RD, Graham KS, Sutton JR, et al. Medium altitude training and sea level performance [abstract]. Med Sci Sports Exerc 1996; 28: S124
26. Nummela ARI. Acclimatization to altitude and normoxic training improve 400-m running performance at sea level. J Sports Sci 2000; 18: 411-9
27. Vallier JM, Chateau P, Guezennec CY. Effects of physical training in a hypobaric chamber on the physical performance of competitive triathletes. Eur J Appl Physiol 1996; 73: 471-8
28. Rodríguez FA, Ventura JL, Casas M, et al. Erythropoietin acute reaction and haematological adaptations to short, intermittent hypobaric hypoxia. Eur J Appl Physiol 2000; 82: 170-7
29. Bailey DM, Davies B, Romer L, et al. Implications of moderate altitude training for sea-level endurance in elite distance runners. Eur J Appl Physiol 1998; 78: 360-8
30. Burtscher M, Nachbauer W, Baumgartl P, et al. Benefits of training at moderate altitude versus sea level training in amateur runners. Eur J Appl Physiol 1996; 74: 558-63
31. Friedmann B, Jost J, Rating T, et al. Effects of iron supplementation on total body hemoglobin during endurance training at moderate altitude. Int J Sports Med 1999; 20: 78-85
32. Gore CJ, Hahn A, Rice A, et al. Altitude training at 2690m does not increase total haemoglobin mass or sea level $\dot{V}O_{2max}$ in world champion track cyclists. J Sci Med Sport 1998; 1: 156-70
33. Ingjer F, Myhre K. Physiological effects of altitude training on elite male cross-country skiers. J Sports Sci 1992; 10: 37-47
34. Jensen K, Nielsen TS, Fikestrand A, et al. High-altitude training does not increase maximal oxygen uptake or work capacity at sea level in rowers. Scand J Med Sci Sports 1993; 3: 256-62
35. Levine BD, Stray-Gundersen J. Altitude training does not improve running performance more than equivalent training near sea level in trained runners [abstract]. Med Sci Sports Exerc 1992; 24: S95
36. Miyashita M, Mutoh Y, Yamamoto Y. Altitude training for improving swimming performance at sea level. Jpn J Phys Fitness Sports Med 1988; 37: 111-6
37. Pyne DB. Performance and physiological changes in highly trained swimmers during altitude training. Coach Sport Sci J 1998; 3: 42-8
38. Rusko HK, Tikkanen H, Paavolainen L, et al. Effect of living in hypoxia and training in normoxia on sea level $\dot{V}O_{2max}$ and red cell mass [abstract]. Med Sci Sports Exerc 1999; 31: S86
39. Saunders PU, Telford RD, Pyne DB, et al. Improved running economy in elite runners after 20 days of simulated moderate-altitude exposure. J Appl Physiol 2004; 96: 931-7
40. Svedenhag J, Saltinj B. Aerobic and anaerobic exercise capacities of elite middle-distance runners after two weeks of training at moderate. Scand J Med Sci Sports 1991; 1: 205-14
41. Svedenhag J, Piehl-Aulin K, Skog C, et al. Increased left ventricular muscle mass after long-term altitude training in athletes. Acta Physiol Scand 1997; 161: 63-70
42. Dehnert C, Huetler M, Liu Y, et al. Erythropoiesis and performance after two weeks of living high and training low in well trained triathletes. Int J Sports Med 2002; 23: 561-6
43. Stray-Gundersen J, Levine BD. Altitude acclimatization normoxic training (high/low) improves sea level endurance immediately on descent from altitude [abstract]. Med Sci Sports Exerc 1994; 26: S64
44. Wehrlin JP, Zuest P, Hallen J, et al. Live high-train low for 24 days increases hemoglobin mass and red cell volume in elite endurance athletes. J Appl Physiol 2006; 100: 1938-45
45. Witkowski S, Karlsen T, Resaland G, et al. Optimal altitude for 'living high–training low' [abstract]. Med Sci Sports Exerc 2001; 33: S292
46. Clark SA, Aughey RJ, Gore CJ, et al. Effects of live high, train low hypoxic exposure on lactate metabolism in trained humans. J Appl Physiol 2004; 96: 517-25
47. Gore CJ, Hahn AG, Aughey RJ, et al. Live high: train low increases muscle buffer capacity and submaximal cycling efficiency. Acta Physiol Scand 2001; 173: 275-86
48. Hahn AG, Gore CJ, Martin DT, et al. An evaluation of the concept of living at moderate altitude and training at sea level. Comp Biochem Physiol A Mol Integr Physiol 2001; 128: 777-89
49. Hinckson EA, Hopkins WG. Changes in running endurance performance following intermittent altitude exposure simulated with tents. Eur J Sport Sci 2005; 5: 15-24
50. Hinckson EA, Hopkins WG, Fleming JS, et al. Sea-level performance in runners using altitude tents: a field study. J Sci Med Sport 2005; 8: 451-7

51. Martin DT, Hahn AG, Lee H, et al. Effects of a 12-day 'live high train low' cycling camp on 4-min and 30-min performance [abstract]. Med Sci Sports Exerc 2002; 34: S157
52. Mattila V, Rusko H. Effect of living high and training low on sea level performance in cyclists [abstract]. Med Sci Sports Exerc 1996; 28: S156
53. Roberts A, Clark S, Townsend N, et al. Changes in performance, maximal oxygen uptake and maximal accumulated oxygen deficit after 5, 10 and 15 days of live high: train low altitude exposure. Eur J Appl Physiol 2003; 88: 390-5
54. Basset FA, Joanisse DR, Boivin F, et al. Effects of short-term normobaric hypoxia on haematology, muscle phenotypes and physical performance in highly trained athletes. Exp Physiol 2006; 91: 391-402
55. Katayama K, Matsuo H, Ishida K, et al. Intermittent hypoxia improves endurance performance and submaximal exercise efficiency. High Alt Med Biol 2003; 4: 291-304
56. Katayama K, Sato K, Matsuo H, et al. Effect of intermittent hypoxia on oxygen uptake during submaximal exercise in endurance athletes. Eur J Appl Physiol 2004; 92: 75-83
57. Gore CJ, Rodriguez FA, Truijens MJ, et al. Increased serum erythropoietin but not red cell production after 4 wk of intermittent hypobaric hypoxia (4,000-5,500 m). J Appl Physiol 2006; 101: 1386-93
58. Rodriguez FA, Truijens MJ, Townsend NE, et al. Performance of runners and swimmers after four weeks of intermittent hypobaric hypoxic exposure plus sea level training. J Appl Physiol 2007; 103: 1523-35
59. Bonetti DL, Hopkins WG, Kilding AE. High-intensity kayak performance after adaptation to intermittent hypoxia. Int J Sports Physiol Perform 2006; 1: 246-60
60. Bonetti DL, Hopkins WG, Kilding AE, et al. Cycling performance following adaptation to two protocols of acutely intermittent hypoxia. 12th Annual Congress of the European College of Sport Science; 2007 Jul 12-15; Jyväskylä
61. Hamlin MJ, Hellmans J. Effect of intermittent normobaric hypoxic exposure at rest on hematological, physiological and performance parameters in multi-sport athletes. J Sports Sci 2007; 25: 431-41
62. Hinckson EA, Hopkins WG, Downey BM, et al. The effect of intermittent hypoxic training via a hypoxic inhaler on physiological and performance measures in rowers: a pilot study. J Sci Med Sport 2006; 9: 177-80
63. Julian C, Gore C, Wilber R, et al. Intermittent normobaric hypoxia does not alter performance or erythropoietic markers in highly trained distance runners. J Appl Physiol 2004; 96: 1800-7
64. Dufour SP, Ponsot E, Zoll J, et al. Exercise training in normobaric hypoxia in endurance runners: I. Improvement in aerobic performance capacity. J Appl Physiol 2006; 100: 1238-48
65. Hendriksen IJ, Meeuwsen T. The effect of intermittent training in hypobaric hypoxia on sea-level exercise: a cross-over study in humans. Eur J Appl Physiol 2003; 88: 396-403
66. Katayama K, Sato Y, Morotome Y, et al. Ventilatory chemosensitive adaptations to intermittent hypoxic exposure with endurance training and detraining. J Appl Physiol 1999; 86: 1805-11
67. Morton JP, Cable NT. The effects of intermittent hypoxic training on aerobic and anaerobic performance. Ergonomics 2005; 48: 1535-46
68. Roels B, Millet GP, Marcoux CJ, et al. Effects of hypoxic interval training on cycling performance. Med Sci Sports Exerc 2005; 37: 138-46
69. Terrados N, Melichna J, Sylvén C, et al. Effects of training at simulated altitude on performance and muscle metabolic capacity in competitive road cyclists. Eur J Appl Physiol 1988; 57: 203-9
70. Truijens MJ, Toussaint HM, Dow J, et al. Effect of high-intensity hypoxic training on sea-level swimming performances. J Appl Physiol 2003; 94: 733-43
71. Ventura N, Hoppeler H, Seiler R, et al. The response of trained athletes to six weeks of endurance training in hypoxia or normoxia. Int J Sports Med 2003; 24: 166-72
72. Hackett PH, Roach R. High-altitude medicine. In: Auerbach PS, editor. Wilderness medicine. 3rd ed. St Louis (MO): Mosby, 1995: 1-37
73. Hopkins WG, Schabort EJ, Hawley JA. Reliability of power in physical performance tests. Sports Med 2001; 31: 211-34
74. Toussaint HM, Hollander AP. Energetics of competitive swimming: implications for training programmes. Sports Med 1994; 18: 384-405
75. Hinckson EA, Hopkins WG. Reliability of time to exhaustion analyzed with critical-power and log-log modeling. Med Sci Sports Exerc 2005; 37: 696-701
76. Batterham AM, Hopkins WG. Making meaningful inferences about magnitudes. Int J Sports Physiol Perform 2006; 1: 50-7
77. Hopkins WG. A spreadsheet for combining outcomes from several subject groups. Sportscience 2006; 10: 51-3
78. Pyne D, Trewin C, Hopkins W. Progression and variability of competitive performance of Olympic swimmers. J Sports Sci 2004; 22: 613-20
79. Hopkins WG. Competitive performance of elite track-and-field athletes: variability and smallest worthwhile enhancements. Sportscience 2005; 9: 17-20
80. Paton CD, Hopkins WG. Variation in performance of elite cyclists from race to race. Eur J Sport Sci 2006; 6: 1-7
81. Schmidt WF, Prommer N, Heinicke K, et al. Impact of total hemoglobin mass on $\dot{V}O_{2max}$ [abstract]. Med Sci Sports Exerc 2007; 39 (5): S3
82. Snowling NJ, Hopkins WG. Effects of different modes of exercise training on glucose control and risk factors for complications in type 2 diabetic patients: a meta-analysis. Diabetes Care 2006; 29: 2518-27
83. Baker A, Hopkins WG. Altitude training for sea-level competition. Sportscience 1998; 1 [online]. Available from URL: http://sportsci.org/traintech/altitude/wgh.html. [Accessed 2009 Jan 18]
84. Rusko HK, Tikkanen HO, Peltonen JE. Altitude and endurance training. J Sports Sci 2004; 22: 928-45

Correspondence: Prof. *Will G. Hopkins*, School of Sport and Recreation, AUT University, Private Bag 92006, Auckland 0627, New Zealand.
E-mail: will@clear.net.nz