

北太平洋宽吻海豚正常生理指标数值的算法*

陈 艳¹ 王文琪² 徐申波² 李 建²

(¹ 四川省西昌师范专科学校生物系 西昌 615000)

(² 莱阳农学院 青岛 266000)

提要 连续 5 年对青岛海豚表演馆中北太平洋宽吻海豚(*Tursiops gillie Dall*)进行现场实验和行为学研究,获取了大量的生理指标数据,运用不同的分析和计算方法对所得数据进行了筛选、剔除及处理,确立了海豚各生理指标的正常阈值和范围确定的方法,从而为准确了解海豚机体的生理状况、疾病诊断及治疗奠定了基础。

关键词 海豚,生理指标,数据算法

中图分类号 Q41 **文献标识码** A **文章编号** 1000 - 3096(2003)12 - 0007 - 03

在海豚的饲养管理和疾病防治工作中,定期体检是一项必须而且重要的工作。由于海豚个体差异很大,很难直接应用体检所得的各项生理指标测定值对其健康状况做出准确的判断。在实践中,通过采取科学的数据处理和算法,初步统计建立北太平洋宽吻海豚正常生理指标的范围,在每头海豚建立其生理指标档案的基础上,便能够准确了解海豚的生理状况。

所运用的数据是通过多头海豚的血常规指标和血液生化指标等进行的 500 余头次的测定(按国际惯例,海洋哺乳动物统计样本数一般为 20 头次^[1])直接获得,包括血色素含量等 9 项血液指标、血糖等 39 项血液生化指标和体温等 45 类原始数据,以下简称具体的数据处理办法和算法。

1 原始数据中超常值的剔除

每头海豚的正常生理指标数值虽有差异,但机体在健康情况下各生理指标值均有一定的范围^[2]。通过数据分析整理,去掉生病海豚的一些超常指标,选择适当的计算方法,便可以归纳出海豚血液指标的正常阈值和范围。设海豚群体为 N 头,任取一海豚个体以 a 表示,则 $1 \leq a \leq N$,其某血液指标的测量值以 x_a 表示,残差以 v_a 表示,标准差以 S 表示,依据拉依达准则,若 v_a 满足 $|v_a| > 3S$ 时,该数值便不在正常范围之内,可以直接剔除该数值(这样的数据通常情况下对应的海豚一般是生病、受伤或其它超常状态)。

2 数据的预处理

2.1 样本的选取

为计算海豚的正常生理指标阈值,首先要选取合适的样本。选择样本需注意以下几点:(1)海豚眼观应处于健康状态,或虽有疾病但对被测算的指标没有影响。(2)选取的样本应有足够的数量,一般在 100 头次以上。(3)所有指标的测定方法前后要统一而准确。测定的仪器、操作规程都要一致,并尽量与今后应用正常值范围的实际情况一致。(4)确定海豚各生理指标取单侧还是双侧阈值,应根据指标的生理学意义的不同而定。一般情况下,各生理指标的正常值范围分别确定其上限和下限的,这是其双侧阈值;而根据实际的生理学意义,某些指标尽有上限或下限的范围,则称为该生理指标的单侧阈值。(5)选定适当的百分界限。确定海豚的各生理指标的正常值范围通常是从正常动物测定结果中取其中间部分。一般以包括 95% 个体测定值的界限作为正常值范围。

* 中国科学院重点基金项目 A35970711 号。

第一作者:陈艳,出生于 1970 年,大学,讲师。电话:0834-2192615, E-mail: ChenYan@163.net

收稿日期:2002-03-22;修回日期:2003-07-18

2.2 样本的分组

由于本研究所分析的数据有 8 000 多个,所以应根据一定的方法进行特定的数据分组处理。以下是具体的分组程序:(1) 求全距:全距即极差,为同项全部数据中最大值与最小值之差。(2) 确定组数:为方便计算,需先将有关的数据按大小次序排列,然后分为若干组。组数应根据样本数的多少和要求的高低而定,分组越多,统计量越精确,但不便于计算,分组的数量一般为:样本数 30~60, 60~100, 100~200, 200~500, 500 以上。组数 6~8, 7~10, 9~12, 12~17, 17~20。(3) 计算组距:每组的范围称为组距,用 I 表示。 $I = \text{全距} / \text{组数}$ 。(4) 选定组限:海豚生理指标的每组数值的起点值为其下限,终点值为其上限,为避免第一组中变量过多,一般第一组的下限比最小值小半个组距,第一组的下限确定后,即可求出各组的组限(后一组的下限就是前一组的上限)。然后将各个观察值按以下规则归入组中:某组下限 $\leq$ 测定值 $<$ 某组上限(详见例子中的频数分布图)。(5) 求组中值:分组以后,每个观察值不再是具有特征的数据,而用所属那一组的组中值为代表,公式为:组中值 = (上限 + 下限) / 2 = 下限 + 1/2 组距。

2.3 分组资料的计算

频数的计算:某组数据的频数,即指归入该组数据中的所有检测值的各数。当数据足够大时,一般用频率作为概率的估计值。某组数据的累计频数,即该组以及之前各组的频数之和,而指标数值的平均数的计算是通过加权法和等级差法取得的。

3 估计正常范围的方法

3.1 百分位数法

求海豚各生化指标的正常值的 $X(\%)$ 的界限,以 P_x 表示。首先将全部检测的数值由小到大依次排列、分组,求出各组的频数 P_x 、累计频数,求指标的双侧阈值时,需分别求出 P_x 和 $P_{(100-x)}$;求指标的单侧阈值时,只需求出 P_x 即可,然后找出累计频数略大于 $N \times X\%$ 的组即可,公式如下:

$P_x = L + I(N \times X - C) / f_x$,式中, L 为 P_x 所在组的下限, I 为组距, f_x 为 P_x 所在组的频数, C 为小于 L 各组的累计频数。

3.2 正态分布法

对于服从正态分布的数据资料,可以应用正态分布原理估计正常范围。通过对所测得的海豚各生理指标数据的分组、计算频数、编秩、正态性 D 检验和估计正常范围等一系列步骤进行分析计算,得出正常范围常用的双侧和单侧百分数表。

注意编秩时,某组的首秩次为上一组的秩次加 1,末秩次为该组的累计频数。第一组的秩次范围为该组的频数。某组的平均秩次为该组的首末秩次之和除以 2。

3.3 数据转换法

如果资料不符合正态分布,但数据通过转换后服从正态分布的,仍可用上述方法估计正常值。数据转换方法常用的有平均根法,对数转换法等。例如采用对数转换的方法如下:(1) 将检测数值中的最大值和最小值转化成对数,求其全距。(2) 先按对数值确定组界,即确定对数组;然后将各组界分别取反对数,得数即为真数组的组界。(3) 按真数组的组界对数据进行分组,计算各组的频数。因真数组的组界与对数组的组距一一对应,以下的计算按真数组进行。(4) 如此计算出的正常值范围为对数值,取其反对数即为正常值范围。

以海豚的羟丁酸脱氢酶指标 (HDBH) 为例,将 5 年多的体检数据进行处理,得出了海豚羟丁酸脱氢酶数值的平均值为 351.94 U/L,其正常上限为 491.856 U/L,正常下限为 212.028 U/L(该指标的所有数据的分布频率如表 1)。利用该指标的正常阈值和体检值的比较,进行海豚肝脏、心脏疾病状况的诊断与治疗等。

表 1 羟丁酸脱氢酶数据的接收范围频率表

Tab.1 The frequency of HDBH data reception

HDBH 数据的接收范围	频率	累积(%)
< 214	1	0.83
$214 \leq X < 245.9091$	3	3.33
$245.9091 \leq X < 277.8182$	11	12.50
$277.8182 \leq X < 309.7273$	19	28.33
$309.7273 \leq X < 341.6364$	26	50.00
$341.6364 \leq X < 373.5455$	22	68.33
$373.5455 \leq X < 405.4545$	11	77.50
$405.4545 \leq X < 437.3636$	13	88.33
$437.3636 \leq X < 469.2727$	6	93.33
$469.2727 \leq X < 501.1818$	3	95.83
$501.1818 \leq X < 533.0909$	2	97.50
其它	3	100.00

又如海豚的乳酸脱氢酶 (LDH) 指标,其数据的分布频率如表 2 所示,利用该指标的正常阈值和体检值的比较,进行海豚肾脏、肝脏及肺等器官疾病的诊断等。

海豚各指标的正常范围主要是运用以上百分位

表 2 乳酸脱氢酶(LDH)数据的接受范围频率表

Tab.2 The frequency of LDH data reception

LDH 数据的接收范围	频率	累积(%)
< 225	1	1.22
$225 \leq X < 273.444\ 4$	3	4.88
$273.444\ 4 \leq X < 321.888\ 9$	4	9.76
$321.888\ 9 \leq X < 370.333\ 3$	11	23.17
$370.333\ 3 \leq X < 418.777\ 8$	25	53.66
$418.777\ 8 \leq X < 467.222\ 2$	12	68.29
$467.222\ 2 \leq X < 515.666\ 7$	9	79.27
$515.666\ 7 \leq X < 564.111\ 1$	9	90.24
$564.111\ 1 \leq X < 612.555\ 6$	6	97.56
其它	2	100.00

数法和正态分布法算出,在此基础上,通过进行每次

体检所得指标值与正常阈值的比较与分析,便能够准确判断海豚的生理状况^[4]。

参考文献

- 1 徐汉光,贾树林,李忠学,等. 黄海北部 F 鳐鲸的研究. 动物学报,1983,29(1):86-92
- 2 Pilleri G, Chen P X. The brain of the Chinese finless porpoise *Neophocaena asiaeorientalis* (Pilleri and Gihl, 1972). I. Macroscopic anatomy. Invest Cetacea, 1982, 13:27-78
- 3 Uchida Y. Several properties of the blood of finless porpoise, *Neophocaena phocaenoides*, in captivity. J of Jap Zoo and Aquarium, 1987, 29(1):17-23
- 4 Chen P X, Liu R J. Captive husbandry of the Baiji, *Lipotes vexillifer*. Occasional Papers of the IUCN Species Survival Commission (SSC), 1989(3):81-85

SCIENTIFIC STATISTICAL METHODS TO ANALYZE THE PHYSIOLOGICAL DATA OF *Tursiops gille* Dall

CHEN Yan¹ WANG Wen - Qi² XU Shen - Bo² LI Jian²

(¹ Biological Department of Xichang Teacher's School, Xichang, 615000)

(² Laiyang Agricultural College, Qingdao, 266000)

Received: Mar., 22, 2002

Key Words: Dolphin, Physiological indicators, Statistical methods of data

Abstract

Based on laboratory research from a number of dolphins (*Tursiops gille* Dall) held in the Qingdao Dolphin Show, more than 5 - years worth of experimental data concerning behavior, health and husbandry of dolphins were collected. Through scientific statistical methods such as selection, elimination and other treatment, a method to set the normal value range of different physiological indicators of dolphins were established. That was a very useful indication of the physical condition, diagnosis and treatment of dolphins.

(本文编辑:刘珊珊)