ABSTRACT

There is currently a growing concern over the use of a representative survey, which could be used to develop standards in physical education and sport. The only reference data that can be usually found is that remaining from the 1980s. One of the new incentives could be the customized adjustment of standards, so that a person can compare themselves with the same group of people, not only by age and sex, as applies to most current traditional standards, but so that the group matches in other parameters, in other words, so that the standards are more differentiated. The aim of our research was to explore the differentiation of physical fitness standards in the Czech population, which in the future will be based on data management and analysis via a knowledge base. This system allows data differentiation according to specified criteria and thus the standard will become dynamic, i.e. relevantly adapted to each user. For our initial analysis, we chose the results of the motor test for the standing long jump in the male population aged 18–19 years, who are already adults, and for whom we have the largest database. Besides differentiation based on sex and age, we have also selected additional criteria for body height, body weight, level of physical activity and the region where the tested person works or attends school. The effect of these parameters that has been monitored based on the analysis of variance. Statistically significant differences occurred with the factors for body height (p-value < 0.0001), body weight (p-value < 0.0001) and the level of physical activity (p-value < 0.0001), while on the other hand, there were no significant statistical differences between groups in the regional factor (p = 0.1458). When evaluating the effect size of the variance analysis through the η^2 coefficient, the regional factor shows very little effect (η^2 = 0.0258), the body height factor and the level of physical activity show a medium effect (η^2 = 0.0715 and η^2 = 0.0775) and the body weight factor shows a great effect (η^2 = 0.1473). Based on these results we consider expanding the number of criteria for the differentiation of standards as effective and appropriate. The analysis of the sorting factors is therefore one of the constituents, which will lead to the successful creation of a knowledge base.

Keywords: analysis of variance; standards; physical fitness; population testing; motor tests

DOI: 10.14712/23366052.2015.27
INTRODUCTION

The Czech Republic previously had a unique tradition of large-scale population surveys, which were particularly developed in the last century, post World War II. The most famous test system was undoubtedly the Tyršův fitness badge, which with the onset of the totalitarian regime in the 1950s was renamed in the USSR style as PPOV and BPPOV (Prepared to work and defend the motherland and Be prepared to work and defend the motherland) (ČSTV, 1975; Novotný, 1961). Test batteries designed in modern way were developed in this country in the early 1970s and were applied in two large-scale nationwide school in 1966 and university in 1965 surveys of young people. The authors of the papers published later are Pávek (1977) and Měkota and Šorm (1972). In the testing of the motor fitness of member of the Czech Union of Physical Education and Sports, which was carried out in two stages in 1972–73 and 1982 and included adults well as senior citizens, a seven items test battery was used. A summarizing report was presented by Kovář (1985).

Nationwide surveys of motor performance were repeated, using modified test batteries, in school population in 1987, in university students in 1986. The test sets are described in the book by Moravec (1990) and in the summarizing paper by Kolář, Měkota and Šorm (1989). In this vast research project were collected data needed for the construction of comprehensive test norms with nationwide validity. Nationwide surveys of motor performance were repeated, using modified test batteries, in school population in 1987, in university students in 1986. The test sets are described in the book by Moravec (1990) and in the summarizing paper by Kolář, Měkota and Šorm (1989). In this vast research project were collected data needed for the construction of comprehensive test norms with nationwide validity. These motor tests were systematically used in physical education classes at Czech universities, serving as diagnostic aids for classification of men and woman students into various categories of physical education classes (Kovář & Měkota, 1995). After the Fall of Communism, this tradition was followed by the test battery – the Unifittest (6–60) (Kovář & Měkota, 1995) and so on (Rychtecký, Tilinger, & Dovalil, 2009; Machačová & Bunc, 2009). It should fill the gap that arose after abolishing the PPOV badge and it should become an integral part of Physical Education lessons (Kovář & Měkota, 1995).

From 1951 up to 2001, nationwide anthropological research was conducted every 10 years, which provided valuable data on secular trends in the Czech population. Unfortunately this tradition was interrupted in 2011 and for the first time since World War II, no research was conducted (SZÚ, 2011). Regarding the post-war period, the Spartakiads (mass gymnastics events) and compulsory military service also provided a considerable opportunity to conduct a population survey. Prior to the last full population survey – the census in 2011, the unwillingness of the population to participate in such a research could clearly be observed. We have therefore lost most of these opportunities to survey the population in a manner that was financially viable with the arrival of democracy and are forced to look for new solutions.

Today’s rapidly changing information and communication technologies provide us with a number of interesting tools that can be used in the collection, processing and analysis of data. The basic principle is the development of a knowledge base, which basic characteristic is the Experience Management (EM), which will gradually prevail over the
preset algorithms (Berka, 2003). A knowledge base enables to continuously and physically create and share information with all subject units and constantly optimize processes to ensure the transfer of information between all participants. The first step in creating a knowledge base is the creation of a basic database, which at the first stage comprises data taken from the Unifittest (6–60), and the results are compared with the differentiated standards according to specified criteria. In order to avoid a large number of criteria filtering and to avoid any unnecessary expansion of the database by data with no real use, we decided to find out if the acquired population data, sorted according to the given criteria in the Czech population, actually differs i.e. if it makes sense to differentiate the standards according to these criteria.

**OBJECTIVE**

The aim of this particular study was to determine whether the male population aged 18–19 years, differ in the motor test results in the standing long jump depending on body height, body weight, the region in which the person works or studies and the level of physical activity and therefore if it is useful to differentiate the standards according to these criteria.

**METHODS**

The sample consisted of 430 men aged 18–19 years (221 eighteen year olds, 209 nineteen year olds) from nine regions in the Czech Republic (1 = Prague, 2 = South Bohemian Region, 3 = South Moravian Region, 4 = Karlovy Vary Region, 6 = Liberec Region, 7 = Moravian-Silesian Region, 8 = Olomouc Region, 9 = Pardubice Region, 10 = Pilsen Region) with a mean body height of 180.3 ± 7.4 cm and a mean body weight of 72.9 ± 10.3 kg. The data from the rest of missing regions were not available, so the selection based on availability (Hendl, 2012) was used.

The sample was not taken at random, but taking into account the size and normal frequency of data distribution from a standing long jump test – see Table 2, where for our testing purposes and the selected methods, the data are considered to be satisfactory.

For basic data analysis, taking into account the size of the contingent table and by combining all the factors, we have used a one-way analysis of variance by testing each factor separately. However, to avoid any interaction of the factors, we also conducted a multi-factor analysis of variance with interactions without repetition. We assumed that each factor level cannot acquire more repeating permutations.

The dependent variable consisted of motor test results from the standing long jump, with selected factors – body height, body weight, the region in which the tested person works or studies and the level of physical activity. For the body height and body weight, intervals of 5 cm and 5 kg were created as ratio variables. The level of physical activity was categorized into nine groups according to the Compass questionnaire (Rychtecký, 2006) (1 = competitive, organized, intense physical activity, where the frequency is greater than 120 times a year; 9 = no sports or physical activity). However, in our data,
only five of these groups are represented. Frequencies in different groups according to the factors are listed in Table 1. The combined factors relative to the size of the contingent table are not given. The difference in each level of the statistically significant factors was not analysed in this first phase; for our purposes, we are satisfied in determining whether at least two factor levels differ, i.e. it will make sense to count on the factor as a filtering criterion. In order to verify the assumption of homoscedasticity of each factor (verification of the hypothesis of equal variances in normal distribution), we did not use the traditional Bartlett’s test, because it is very sensitive to the violation of the assumption of normality and although the normality of our data was not simply rejected, we used the Fligner-Killeen test, which is not so sensitive, but is able to verify the homoscedasticity at a sufficient level (Neubauer, 2011).

In order to verify the results of the variance analysis, we had to further determine the effect size of the variance analysis through the $\eta^2$ coefficient, which meets all its variants (Hayes, 2009). By multiplying the $\eta^2$ coefficient by one hundred, the result is the percentage of the explained variance of the dependent variable of the factor used. According to Cohen (1988) $\eta^2 \geq 0.0099$ means little effect, $\eta^2 \geq 0.0588$ means moderate effect and $\eta^2 \geq 0.1379$ means a great effect.

Data normality was tested using NCSS 2007 software, other analysis was conducted using R Project (R-2.15.3) software.

### Table 1. Frequencies in individual groups classified by factors

| Body weight | No. of PT | Body height | No. of PT | Level of physical activity | No. of PT | CZ Region | No. of PT |
|-------------|-----------|-------------|-----------|---------------------------|-----------|-----------|-----------|
| 50          | 9         | 160         | 12        | 1                         | 32        | 1         | 40        |
| 55          | 21        | 165         | 20        | 2                         | 35        | 2         | 57        |
| 60          | 60        | 170         | 54        | 3                         | 57        | 3         | 25        |
| 65          | 81        | 175         | 102       | 4                         | 44        | 4         | 14        |
| 70          | 94        | 180         | 123       | 9                         | 262       | 6         | 28        |
| 75          | 65        | 185         | 74        |                           | 7         | 69        |           |
| 80          | 46        | 190         | 38        |                           | 8         | 39        |           |
| 85          | 17        | 195         | 7         |                           | 9         | 125       |           |
| 90          | 19        |             |           |                           | 10        | 36        |           |
| 95          | 10        |             |           |                           |           |           |           |
| 100         | 8         |             |           |                           |           |           |           |
| sum         | 430       | sum         | 430       | sum                       | 430       | sum       | 430       |

PT = people tested

### RESULTS

The actual data analysis was preceded by verification of the preconditions for using the analysis of variance.
Table 2. Normality test of frequency distribution

| Normality Test Section of jump | Test Name      | Value     | Prob Level   | 10% Critical Value | 5% Critical Value | Decision 5%          |
|-------------------------------|----------------|-----------|--------------|--------------------|-------------------|----------------------|
|                               | Shapiro-Wilk W | 0.964851  | 0.219599     |                    |                   | Cannot reject normality |
|                               | Kolmogorov-Smirnov | 0.0916   |              | 0.124              | 0.135             | Cannot reject normality |

The normality of the frequency distribution was tested using a wide range of tests offered by NCSS 2007 software; Table 2 shows the results of frequently used Shapiro-Wilk W and Kolmogorov-Smirnov tests. All tests consistently show that normality cannot be rejected.

The homoscedasticity of the monitored factors was tested using the Fligner-Killen test; the results are shown in Table 3.

Table 3. Fligner-Killen homogeneity of variance test

| Fligner-Killeen homogeneity of variance test | chi-squared | df | p-value |
|--------------------------------------------|-------------|----|---------|
| Standing long jump × region                | 12.1586     | 8  | 0.1443  |
| Standing long jump × body height           | 9.2715      | 7  | 0.2337  |
| Standing long jump × body weight           | 9.4172      | 10 | 0.493   |
| Standing long jump × level of physical activity | 8.9788     | 4  | 0.06163 |

All p-values are greater than 0.05 i.e. at 5% level of significance so it can be stated that the variances in the different groups classified by different factors do not differ.

The distribution of motor test results for the standing long jump into groups by factors is shown in the boxplots in Figures 1–4.

![Figure 1. Standing long jump by region](image-url)
Figure 2. Standing long jump by body height

Figure 3. Standing long jump by body weight
For the regional factor in Figure 1, the greatest evenness of the mid values for all groups is apparent. For the body height factor in Figure 2 and the body weight factor in Figure 3, the large intergroup variance is immediately apparent. For the body weight, we can see that the best results are achieved by individuals from the 77–84 kg group; on the other hand, the results plummet for weight over 94 kg.

For the body height it is obvious that the intergroup variance is great in extreme groups, on the other hand for medium size groups, the intergroup variance is small. For the level of the physical activity factor in Figure 4, the individual groups have fairly balanced medium values and variances; the only difference applies to the last group – number 9, which includes individuals with the lowest levels of physical activity.

Table 4 shows the results of a one-way analysis of variance.

| Table 4. Results of a one-factor analysis of variance |
|-----------------------------------------------------|
| Df       | Sum Sq | Mean Sq | F value | Pr (>|F|) |
| Region   |        |         |         |          |
| 8        | 10911   | 1364    | 1.5265  | 0.1458   |
| Residuals| 14      | 376150  | 893     |          |
| Body height | 1    | 17958   | 17958   | 20.824   | <0.0001 *** |
| Residuals | 428    | 369103  | 862     |          |
| Body weight| 10   | 44456   | 4446    | 5.4369   | <0.0001 *** |
| Residuals | 419    | 342605  | 818     |          |

Figure 4. Standing long jump by the level of physical activity
The already apparent finding from the boxplots was once again established, i.e. that statistically significant differences exist among the factors for body height, body weight and the level of physical activity at a significance level of less than 0.001. On the other hand, the statistically significant differences for the factor region between groups were not established.

Table 5. Results of the multi-factor analysis of variance with interactions without repeating

|                          | Df | Sum Sq  | Mean Sq | F value  | Pr (>F) |
|--------------------------|----|---------|---------|----------|---------|
| Level of physical activity | 4  | 28965   | 7241    | 8.5942   | <0.0001 *** |
| Residuals                | 425| 358095  | 843     |          |         |

Pr (>F) = the p-value associated with the F statistic
*** p-value ≤ 0.001

The results of the multi-factor (specifically four-factor) analysis of variance with interactions without repetitions are shown in Table 5. Even with multiple classifications, the individual factors appear to have the same results, i.e. statistically significant
differences between the groups of factors are body height, body weight and the level of physical activity. Compared to the one-way analysis, a 5% significance level also showed differences in the regional factor. Differences which can be attributed to the interaction of factors: body weight – region and body height – body weight – region are also at the same level of significance. For the results of a multi-factor analysis of variance we must take into account the fact that some groups of combined factors were not represented at all and many of these had a low frequency of people tested.

The results of the variance analysis are also in line with the established size of the effect through the η² coefficient. The coefficient values for statistically significant factors are shown in Table 6.

Table 6. Results of the effect size determination through the η² coefficient

| Factors                             | η²     |
|-------------------------------------|--------|
| Region                              | 0.0258 |
| Body height                         | 0.0715 |
| Body weight                         | 0.1472 |
| Level of physical activity          | 0.0775 |
| Body weight – Region                 | 0.1473 |
| Body height – Body weight – Region   | 0.0838 |

The regional factor shows according to Cohen interpretation (1988) very little effect, the body height factor and the level of physical activity show a medium effect and the body weight factor shows a great effect. The combined body height, body weight and regional factors show the medium effect size and the combined body weight and regional factors show a great effect size.

DISCUSSION

The data analysis conducted suggests that the current adult population, although in our research the sample is only represented by the 18–19 years old male category, is on the level of the explosively-power capabilities of the lower extremities differentiated according to three of the four factors selected. We have only analyzed the results of the test – standing long jump out of the entire spectrum of motor tests that are included in the various test systems, because the largest amount of data available to us is from the adult population. Based on the results, we can assume that even in the case of other motor tests, which are included in the battery of tests aimed at physical fitness, the results would be similar. Although numerous research analyses addressing physical fitness and focusing on a specific factor (medical disability, the elderly population, biological age of children, obesity etc.) already exist (Brahler, 2004; Ka Yee Wong, 2006; Miyatake, Miyachi, Tabata, & Numata, 2012; Rikli & Jones, 1999), obtaining population data from the healthy adult population appears to be a significant problem, and not just in the Czech Republic. Baumgartner, Jackson, Mahar, and Rowe (2007) suggested that the sample size
should be several hundred with scores collected over several years. Morrow, Jackson, Disch, and Mood (2005) stated that the sample size should be at least 200 per gender (group). But the sample in the present study unfortunately doesn’t meet the sample size criteria for each group, because our factors have lots of groups and it was impossible for us to obtain 200 people per group. Strand, Hjelm, Shoepe, and Fajardo (2014) also differed standards according physical activity factor in their study of individuals aged 19 to 20 and like us they came to the following conclusion: arithmetic differences were seen for all levels of physical activity (e.g. never, rarely, 1–2 times/week, etc.). The study of Condon and Cremin (2013) examined relationships between height, gender, weight and lower limb muscle power to balance performance, but correlations were in most groups nonsignificant. On the other hand balance test evaluating coordination abilities depend on other factors than our explosively-power capabilities.

Published standards of widely used Unifittest 6–60 test batteries (Kovář & Měkota, 1995), Eurofit (Pekka & Tuxworth, 1997) and Fitnessgram (Suchomel, 2003) for one thing, no longer correspond to the current population and are only differentiated by age and sex, which is not suitable for the sporting population, which is always compared with those standards as above average nor the handicapped population, which in different ways is then below average. This is confirmed by other studies. Baker, Heath, Smith, and Oden (2011), Gjonbalaj, Gllareva, Gjinovci, and Miftari (2015) and of Purashwani, Datta, and Purashwani (2010) have shown that we cannot compare physically active population (students of sports schools and table tennis players) with the standards of the general population and it is necessary to draw up special norms for these groups. In accordance with our findings turns out that especially the level of physical activity is very important factor.

The issue of standards of test batteries was, for example, addressed by Sharon Plowman (1992), who compared the standards of the most frequently used test batteries in the USA and concluded that the percentile standards, which may have a very different rating of the same performance under different test batteries are particularly problematic, and that they must be revised. However, the creation of updated plus differentiated standards, poses substantial challenges to the amount of data required and to the effective management and analysis of this data. The current trend in this regard are information databases, and which have been established in many fields for the purpose of managing and administering data analysis (Adams, 2010). The analysis of the sorting factors presented in this article is therefore one of the constituents, which, we hope, will lead to the successful creation of a knowledge base.

CONCLUSION

We have demonstrated in our constituent study, based on the results of the analysis of variance of the motor test in the standing long jump, that the male population aged 18–19 years shows different results depending on body height, body weight and the level of physical activity and it is therefore advisable to differentiate according to these criteria, in particular for the newly designed standards. We were unable to demonstrate to a sufficient degree, the differences in results depending on the region in which the tested person works or studies.
The current standards can well rank only mean individuals. In practice, differentiated standards make a better comparison possible, because person can compare himself with similar individual, like himself.

ACKNOWLEDGEMENTS

The paper was written with the institutional support PRVOUK P39 – Social Scientific Aspects of Human Movement Studies.

REFERENCES

Adams, T. B. (2010). DataBase: Research and Evaluation Results. American Journal of Health Promotion, 25(2), 147–151.

Baker, U. C., Heath, E. M., Smith, D. R., Oden, G. L. (2011). Development of Wingate Anaerobic Test Norms for Highly-Trained Women. Journal of Exercise Physiology Online, 14(2), 68–79.

Baumgartner, T. A., Jackson, A. S., Mahar, M. T., Rowe, D. A. (2007). Measurement for evaluation in physical education and exercise science. Boston: McGraw-Hill.

Berka, P. (2003). Dobývání znalostí z databází [Knowledge Discovery in Databases]. Prague: Academia.

Brahler, C. J. (2004). QTc Is Associated with Obesity and Physical Activity Level for African American Females. Journal of Exercise Physiology Online, 7(2), 16–24.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hove: Lawrence Erlbaum Associates.

Condon, C., Cremin, K. (2014). Static Balance Norms in Children. Physiotherapy Research International, 19(1), 1–7.

ČSTV [Czechoslovak Union of Physical Education and Sport] (1975). Metodické pokyny k plnění ideově výchovné části Odznaku zdatnosti Připraven k práci a obraně vlasti (BPPOV a PPOV) [Methodical guidelines to perform ideological education part of Fitness badge Ready for work and homeland defense]. Prague: Olympia.

Gjonbalaj, M., Gllareva, I., Gjinovci, B., Miftari, F. (2015). The status of students of the Faculty of Physical Education and Sport in comparison with standard parameters of the Illinois agility test. Sport Mont, 13(43–45), 225–228.

Hayes, A. F. (2009). Statistical Methods for Communication Science. New Jersey: Lawrence Erlbaum Associates.

Hendl, J. (2012). Přehled statistických metod. Analýza a metaanalýza dat [Summary of statistical methods. Data analysis and meta-analysis]. Prague: Portál.

Ka Yee Wong, A. (2006). Gross Motor Skills Performance of Hong Kong Chinese Children. Journal of Physical Education & Recreation, 12(2), 23–29.

Kolář, V., Měkota, K., Šorm, G. (1989). Pohybová výkonnost a tělesný rozvoj studujících 1. ročníku vysokých škol ČSSR 1986 [Motor performance and physical development of students in the first year of the Czechoslovak university 1986]. In: Sborník Tělesná kultura. Olomouc: Palacký University Olomouc.

Kovář, R. (1985). Testy a normy základní pohybové výkonnosti [Tests and standards of basic motor performance]. Prague: ČÚV ČSTV.

Kovář, R., Měkota, K., et al. (1995). Unifittest (6–60): test and norms of motor performance and physical fitness in youth and in adult age. Olomouc: Palacký University Olomouc.

Macháčková, K., Bunc, V. (2009). Methods of functional status evaluation and available instruments appropriate for older adults: A systematic Review. AUC Kinanthropologica, 45(2), 43–54.

Měkota, K., Šorm, G. (1972). Physical efficiency of Czechoslovak students in their first year at universities and colleges. Tělovýchovný sborník, supplementum 2. Prague: SPN.

Miyatake, N., Miyachi, M., Tabata, I., Numata, T. (2012). Evaluation of anthropometric parameters and physical fitness in elderly Japanese. Environmental Health & Preventive Medicine, 17(1), 62–68.
Moravec, R. (1990). *Telesný, funkčný rozvoj a pohybová výkonnost 7–18-ročnej mládeže v ČSFR* [Physical, functional development and motor performance of 7–18 years old youth from ČSFR]. Bratislava: Šport.

Morrow, J. R., Jackson, A. W., Disch, J. D., Mood, D. P. (2005). *Measurement and evaluation in human performance*. Champaign, IL: Human Kinetics.

Mozumdar, A., Liguori, G., Baumgartner, T. A. (2010). Additional Revised Push-Up Test Norms for College Students. *Measurement in Physical Education & Exercise Science*, 14(1), 61–66.

Neubauer, J. (2011). *Analýza rozptylu*. Prednáška Statistika II – Ekonometrie [Analysis of Variance. Lecture of Statistics II – Econometry]. Retrieved from http://k101.unob.cz/~neubauer/pdf/analyza_rozptylu.pdf.

Novotný, J. (1961). *Kdy, kde, jak soutěžit o odznak zdatnosti. Připraven k práci a obraně vlasti* [When, where, how to compete on fitness badge Ready for work and defense of the homeland]. Prague: STN.

Pávek, F. (1977). *Tělesná výkonnost 7–19leté mládeže ČSSR* [Physical performance of 7–19 years old youth from ČSSR]. Prague: SPN.

Plowman, S. A. (1992). Criterion Referenced Standards for Neuromuscular Physical Fitness Tests: An Analysis. *Pediatric Exercise Science*, 4(1), 10–19.

Purashwani, P., Datta, A. K., Purashwani, M. (2010). Construction of Norms for Skill Test Table Tennis Players. *International Journal of Table Tennis Sciences*, 11(6), 93–98.

Rikli, R. E., Jones, C. J. (1999). Functional fitness normative scores for community-residing older adults, ages 60–94. *Journal of Aging & Physical Activity*, 7(2), 162–181.

Rychtecký, A. (2006). *Monitorování účasti mládeže ve sportu a pohybové aktivitě v České republice* [Monitoring of youth participation in sport and physical activity in the Czech Republic]. Prague: Charles University in Prague.

Rychtecký, A., Tilinger, P., Dovalil, J. (2009). Physical performance related to the participation in sport and physical activities and changes in the last thirteen years in Czech youth. *AUC Kinanthropologica*, 45(2), 15–32.

Strand, S. L., Hjelm, J., Shoeppe, T. C., Fajardo, M. A. (2014). Norms for an Isometric Muscle Endurance Test. *Journal of Human Kinetics*, 40(4), 93–102.

Suchomel, A. (2003). Současné přístupy k hodnocení tělesné zdatnosti u dětí a mládeže (FITNESSGRAM) [Current approaches to evaluation of physical fitness among children and youth (FITNESSGRAM)]. Česká kinantropologie, 7(1), 81–89.

SZÚ (Producer). (2011). *Celostatní antropologické výzkumy (CAV)* [National anthropological researches (CAV)]. Státní zdravotní ústav. Retrieved from http://www.szu.cz/publikace/data/celostatni-antropologicke-vyzkumy-cav.

Šárka Honsová
honsova@ftvs.cuni.cz