ARE SMALL (THOUGH STATISTICALLY SIGNIFICANT) MEAN DIFFERENCES PHYSIOLOGICALLY MEANINGFUL?

To the Editor:

The study by Wagner et al. (30) that compared measurements by the BOD POD air displacement plethysmograph (BP), hydrostatic weighing (HW), and dual energy x-ray absorptiometry (DXA) deserves comment. We strongly disagree with the authors’ suggestion that HW is more accurate than BP because HW provided measurements that were more similar to DXA. Researchers disagree about whether DXA should be considered a reference method for soft tissue measurement (10,15,23,28), in part because DXA results differ widely between (8,13,27), and within manufacturers (5,20,25), and can depend upon tissue thickness (11,16). We recommend that future BP validation studies include alternative reference methods when possible, such as multicompartiment models based on body density, water, and bone mineral content (BMC) (such as studies by Collins et al. (2) and Fields and Goran (6)), or elemental analysis from in vivo neutron activation. This stated, it is important to emphasize that an inherent limitation of nearly all body composition “validation” studies is the lack of a true reference method, because no currently available technique can be regarded as the absolute standard (short of chemical carcass analysis).

Wagner et al. (30) largely base their conclusion that the BP is inaccurate on the “small but significant” mean difference of 1.6%fat between BP and DXA, even though other statistical criteria for good agreement between the two methods were met. Furthermore, in contrast to their findings, other adult studies have shown that BP measured slightly lower than DXA (−0.1 to −3.0%fat) (2,17,19,24), and one had mixed results (14) (at least one of these (24) was published at the time of submission but not cited by Wagner et al.). Moreover, studies comparing BP and HW (including Wagner et al. (30)) have reported mean differences ranging from −4.0 to 2.2%fat (1,2,4,7,9,17,30,31), including two that found nearly identical means (18,19).

Within-day coefficients of variation (CV%) for repeated measurements for these methods have varied between 2–4%fat for both BP and HW in the same individuals (9,18), and 2–4%fat for DXA (5,10,26). Thus, the mean differences between methods reported by Wagner et al. (30) and others (1,2,4,7,9,14,17–19,24,31) are clearly within measurement error. Besides measurement error, inconsistent findings among studies may further be attributed to varying laboratory techniques and equipment. For example, different laboratories have different HW systems and techniques, including various methods for measuring residual lung volume (e.g., helium dilution, oxygen dilution, nitrogen washout), different weighing devices in water and on land, different protocols, etc. Similarly, DXA results can differ by manufacturer and software version (29). The possibility that different BOD PODs may also vary cannot be disregarded. Without a multisite comparison study, it is difficult to ascertain specifically how interlaboratory differences might have contributed to the discrepant findings among studies. Given the methodological variation among laboratories, differing subject populations, and expected within-method repeat abilities documented above, the physiological importance of relatively small mean differences between methods should be carefully contemplated, regardless of statistical significance.

One other point we wish to address is the measurement of thoracic gas volume ($V_{TG}$) by the BP. Wagner et al. (30) suggest that BP-$V_{TG}$ measurements may be too high and that one way to check this would be for the manufacturer to reprogram the BP so that $V_{TG}$ could be measured at the end of a normal exhalation (i.e., at functional residual capacity, FRC) rather than at mid-exhalation, which would allow for a comparison between BP-$V_{TG}$ (if at FRC) and dilution-FRC. In fact, plethysmographic- and dilution-measured FRC would not necessarily agree because dilution gases would not exchange with any trapped gas within the thorax or abdomen, resulting in an underestimate of FRC (12,22). For this reason, plethysmographic determination of $V_{TG}$ should be regarded as the method of choice for measuring lung volumes in healthy and diseased individuals (3,12,21). We suggest that one alternative method for validating BP-$V_{TG}$ would be to compare it with $V_{TG}$ measured by standard pulmonary plethysmography with corrections for tidal volume.

Finally, we agree with the conclusion by Wagner et al. (30) that more research is necessary before the BOD POD is accepted as a densitometric reference method; however, we do not agree that a 1.6%fat significant difference between methods is physiologically meaningful. Given that there are interindividual and interstudy differences between methods, we hope that future research will be directed at investigating underlying reasons for these differences so that appropriate steps to-
ward improving the BOD POD as well as other methods can be made.

Megan A. McCrory, Ph.D.
Energy Metabolism Laboratory
Jean Mayer USDA Human Nutrition Research Center on Aging at Tufts University
David A. Fields, Ph.D.
Division of Physiology and Metabolism
Dept. of Nutritional Sciences
University of Alabama at Birmingham
Paul A. Molé, Ph.D., FACSM
Human Performance Laboratory
University of California at Davis

REFERENCES


27. Tothill, P., A. Avenell, J. Love, and D. M. Reid. Comparisons between Hologic, Lunar and Norland dual-energy X-ray


RESPONSE

We agree with many of the comments provided by McCrory et al., such as there is a “lack of a true reference method” and “no currently available technique can be regarded as the absolute standard.” However, we stand by our decision to use dual-energy X-ray absorptiometry (DXA) as the reference measure in our study (12) and disagree with their recommendation that “future Bod Pod (BP) validation studies include alternative reference methods when possible, such as multicomponent (4-C) models based on body density, water, and bone mineral content (BMC)”. Although 4-C models are generally recognized as the preferred reference method in body composition research, its use is not appropriate in validating the BP. Because a multicomponent model includes a measure of body density (Db), it will be biased toward whichever method—BP or hydrostatic weighing (HW)—that was used to obtain the Db measurement included in the 4-C formula.

Although there continues to be some debate over the use of DXA as a reference measure for estimating relative body fat (%BF), investigators comparing different reference methods to 4-C models have found DXA to be a better predictor of mean %BF than HW, total body water, or potassium-40 measurements (3,4,9,11). Three of these studies reported mean differences in %BF ranging from only −0.4% to 0.4% between DXA and 4-C models (3,9,11), whereas Fuller et al. (4) reported a 1.3% difference that was still not significant. Given that a) a 4-C model would be biased toward the method used to obtain Db, b) DXA provides %BF estimates independent of body density, and c) past research supports the use of DXA as an alternative reference measure, we believe that DXA was the most appropriate reference method for this study.

Numerous recent studies using the BP have been published in the time span between the original submission of our manuscript (12) and its publication. McCrory et al. stated that there were small mean differences in %BF between BP and DXA in these investigations. However, it should be noted that in at least four studies (1,5,6,10) small differences in %BF between BP and DXA were “statistically significant”. Furthermore, two of these studies (1,6) reported good agreement between HW and DXA, but “statistically significant” differences between BP and DXA—just like our study (12).

We do not dispute that the BP met most of our validation criteria, and the mean difference in %BF of 1.6% between BP and DXA was small. Are small, but “statistically significant” differences “physiologically significant”? That question may best be answered by the readers (researchers) based on the level of accuracy that they desire in their work. Should there be more stringent validation criteria for a “reference method” (the intended use of the BP) than a “field method?” McCrory et al. stated that the mean differences in %BF in studies comparing BP and HW ranged from −4.0% to 2.2%. It should be noted that theorized total error (biological plus technical) of %BF using the common “field method” of measuring skin-fold thickness is 3.3% (7).

The BP is a wonderful addition to the field of body composition. Due to the rapidity and ease of testing for both the subject and technician, and the low maintenance of the BP, we believe it can be a valuable tool to “estimate” %BF. Additionally, there is some evidence that it has better reliability than HW (2,8). But multiple reports of “statistically significant” differences between the BP and other methods raise concern about its acceptability as a “reference” method. Ultimately, we stand by our decision to use DXA as the reference measure in this particular study, and our findings that 1) the BP and HW were significantly different and 2) there was better agreement between HW and DXA than between the BP and DXA (12). Our conclusion that more research is needed before the BP can be accepted as a “reference” method is also supported by McCrory et al.

Dale R. Wagner, Ph.D., CSCS
Assist. Prof. of Exercise Science, Vanguard University of Southern California
Vivian H. Heyward, Ph.D., FACSM
Prof. Emeritus, University of New Mexico
Ann L. Gibson, Ph.D.
Assist. Prof. of Exercise Science, University of Akron

REFERENCES


STATEMENT OF OWNERSHIP AND CIRCULATION (Required by 39 U.S.C. 3685)

1. Publication Title: MEDICINE AND SCIENCE IN SPORTS AND EXERCISE
2. Publication no.: 0195-9131. 3. Filing date: 10-01-03. 4. Frequency of issue: Monthly
5. No. of issues published annually: 12. 6. Annual subscription price: $211.00
7. Complete mailing address of known office of publication: Lippincott Williams & Wilkins, 16522 Hunters Green Parkway, Hagerstown, MD 21740. 8. Complete mailing address of the headquarters of general business offices of the publisher: 530 Walnut Street, Philadelphia, PA 19106
9. Full names and complete mailing address of publisher, editor, and managing editor: Publisher: Lippincott Williams & Wilkins, Inc. 530 Walnut Street, Philadelphia, PA 19106; Editor: Peter B. Raven, Ph. D., University of North Texas, Health Science Center at Fort Worth, 3500 Camp Bowie Blvd, Fort Worth TX 76107. Managing Editor: Alisa Perel, MSSE Editorial Office, 410 W Michigan St, Indianapolis IN 46202. 10. Owner: American College of Sports Medicine, c/o D. Mark Robertson, Assistant Executive Vice President, P O Box 1440, Indianapolis IN 46202-1440, 401 West Michigan Street, Indianapolis IN 46202-3232. 11. Known bondholders, mortgagees, and other security holders owning or holding 1% or more of total amount of bonds, mortgages, or other securities: None.

12. Purpose, function, and nonprofit status: Has not changed during preceding 12 months.
13. Publication Name: MEDICINE AND SCIENCE IN SPORTS AND EXERCISE
14. Issue Date for circulation data: Volume 32, Issue 8
15. Extent and nature of circulation: Average number of copies each issue during preceding 12 months: (a) Total No. copies (Net Press Run), 22,000. (b) Paid and/or requested circulation: (1) Paid/requested outside-county mail subscriptions stated on form 3541. (Include advertiser's proof and exchange copies), 15,261. (2) Paid in-county subscriptions (include advertiser's proof and exchange copies), N/A. (3) Sales through dealers and carriers, street vendors, counter sales, and other non-USPS paid distribution, N/A. (4) Other classes mailed through the USPS, N/A. (c) Total paid and/or requested circulation (sum of 15b (1), (2), (3), and (4)), 15,261. (d) Free distribution by mail (samples, complimentary, and other free), (1) Outside-county as stated on form 3541, 170, (2) In-country as stated on form 3541, N/A. (3) Other classes mailed through the USPS, N/A. (e) Free distribution outside the mail (carriers or other means), N/A. (f) Total free distribution (sum of 15d and 15e), 170. (g) Total distribution (Sum of 15c and 15f) 15,431. (h) Copies not distributed, 6569 (i) Total (sum of 15g and h), 22,000. (j) Percent paid and/or requested circulation (15c divided by 15g times 100) 99%, No. copies of single issue published nearest to filing date: (a) Total no. copies (Net Press Run), 21,500. (b) Paid and/or requested circulation: (1) Paid/requested outside-county mail subscriptions stated on form 3541. (Include advertiser's proof and exchange copies), 15,939. (2) Paid in-county subscriptions (include advertiser’s proof and exchange copies), N/A. (3) Sales through dealers and carriers, street vendors, counter sales, and other non-USPS paid distribution, N/A. (4) Other classes mailed through the USPS, N/A. (c) Total paid and/or requested circulation (sum of 15b (1), (2), (3), and (4)), 15,939. (d) Free distribution by mail (samples, complimentary, and other free), (1) Outside-county as stated on form 3541, 180, (2) In-county as stated on form 3541, N/A. (3) Other classes mailed through the USPS, N/A. (e) Free distribution outside the mail (carriers or other means), N/A. (f) Total free distribution (sum of 15d and 15e), 180. (g) Total distribution (Sum of 15c and 15f), 16,119. (h) Copies not distributed, 5,381. (i) Total (sum of 15g and h), 21,500. (j) Percent paid and/or requested circulation (15c divided by 15g X 100), 99%, 15c divided by 15g X 100,

16. This Statement of Ownership will be printed in volume 32, issue 12 of this publication.
17. I certify that the statements made by me above are correct and complete.

Michael Wells
Manager, Periodical Operations