"ความน่าเชื่อถือ ความถูกต้อง ความแม่นยำ และ ความเที่ยงตรง" ความสอดคล้องในวิธีการและความคลาดเคลื่อนจากการวัด ของการวิจัยทางด้านสรีรวิทยาการออกกำลังกาย

"Reliability, Validity, Accuracy and Precision on Exercise Physiology Research" Method Agreement and Measurement Error in the Physiology of Exercise

Received: September 7, 2018
Revised: November 19, 2018
Accepted: November 22, 2018

กิตติคุณ แสงนิล (Kittikun Sangnin)*
ประสพชัย พสุนนท์ (Prasopchai Pasunon)**

บทคัดย่อ

วิธีการวัดที่ถูกต้องและเหมาะสม สามารถช่วยลดความคลาดเคลื่อนที่อาจเกิดขึ้นได้จากกระบวนการออกแบบของการวิจัย บทความนี้ให้ความสำคัญกับการศึกษาถึงปัจจัยที่มีอิทธิพลต่อคุณภาพของการวิจัยทางด้านสรีรวิทยาของการออกกำลังกาย มีปัจจัยที่สำคัญคือ ความน่าเชื่อถือ ความถูกต้อง ความแม่นยำ และความเที่ยงตรง ซึ่งจากปัจจัยข้างต้นนั้นส่งผลให้ผู้วิจัยทางด้านสรีรวิทยาการออกกำลังกาย ต้องพิจารณาถึงประเด็นแนวทางการเลือกใช้เครื่องมือวัด โดยต้องอาศัยเกณฑ์ในการประเมินเครื่องมือวัดที่เหมาะสม โดยมีประเด็นที่สำคัญที่นักวิจัยต้องมุ่งเน้นพิจารณาคือ ความจำเพาะของการวัด ความสามารถในการวัดซ้ำ ความถูกต้องของการวัดหรือวิธีการวัด ความไวของการวัดหรือเครื่องมือการวัด ยิ่งไปกว่านั้น ผู้วิจัยควรดำเนินการตรวจสอบเพื่อลดความคลาดเคลื่อนที่อาจเกิดขึ้นจากกระบวนการวัด ซึ่งเป็นวิธีการที่สำคัญหนึ่งในกระบวนการวิจัยที่ผู้วิจัยจำเป็นต้องพิจารณาตรวจสอบและประเมินในขั้นตอนของการออกแบบการวิจัย โดยแนวทางการตรวจสอบที่ต้องพิจารณา ประกอบด้วย การกำหนดขอบเขตของการวิจัย การตรวจสอบความคลาดเคลื่อนเชิงระบบ การตรวจสอบความคลาดเคลื่อนแบบไม่แน่นอน และความแม่นยำทางสถิติ ดังนั้น ทั้งการพิจารณาประเด็นที่สำคัญและแนวทางเหล่านี้ที่กล่าวมาข้างต้น สามารถทำให้งานวิจัยทางสรีรวิทยาของการออกกำลังกายมีคุณภาพมากยิ่งขึ้น เพื่อเป็นประโยชน์ในการวิจัยและการนำไปประยุกต์ใช้ของหลักการและวิธีการตรวจสอบและประเมินการวัดได้อย่างเหมาะสมต่อนักวิทยาศาสตร์การกีฬา นักสรีรวิทยา โค้ชหรือนักกีฬา

คำสำคัญ : ความสอดคล้องในวิธีการ ความคลาดเคลื่อนจากการวัด ความน่าเชื่อถือ ความถูกต้อง ความแม่นยำ ความเที่ยงตรง

* กิตติคุณ แสงนิล อาจารย์ประจำ คณะวิทยาศาสตร์การกีฬา มหาวิทยาลัยเกษมบัณฑิต
Lecturer in Faculty of Sport Science, Kasem Bundit University, kittikun.san@kbu.ac.th, 095-8846464
** รศ.ดร. ประสพชัย พสุนนท์ รศ.ดร. ประจำคณะวิทยาการจัดการ, มหาวิทยาลัยศิลปากร
Associate Professor Dr., Faculty of Management Science, Silpakorn University, prasopchai@ms.su.ac.th, 032-594-043 (41399)

Abstract

Accurate and appropriate measurement can reduce the errors that may occur from the design process of the research. This article focuses on the study of factors influencing the quality of physiological research of exercise. There are important factors; included reliability, validity, accuracy and precision. From the above factors, the exercises physiologist should consider the selection issue of measurement tool; it must be based on the criteria for measuring appropriate measurement. There are important issues that researchers need to focus on; included specificity, repeatability, validity of measurement methods, and sensitivity. Moreover, researchers should investigate to reduce errors, or errors that may occur from the measurement process; this is one of the important approaches in the research process. The researcher needs to consider, check and evaluate at the stage of research design. The examination guidelines to be considered include delimitations, systematic error examination, random error examination and statistical precision. Therefore, considering both these important issues and these guidelines will make the physiological research of exercise more quality; this would be useful in the research and application of the principles and methods of monitoring and evaluating measurements appropriately for sports scientists, physiologists, coaches or athletes.

Keywords: Method Agreement, Measurement Error, Reliability, Validity, Accuracy, Precision

Introduction/บทนำ

การพัฒนาความรู้ความเข้าใจเกี่ยวกับความสามารถในการออกกำลังกายของมนุษย์ ผลในทางปฏิบัติจากการศึกษาหรือการวิจัยทางด้านสรีรวิทยาของการออกกำลังกายคือ การเพิ่มประสิทธิภาพและความทนทานต่อการออกกำลังกายของบุคคลที่ผ่านการทดสอบ ซึ่งควรจำเป็นต้องมีกระบวนการประเมินก่อนในความต้องการในการวิเคราะห์แบบเต็มรูปแบบ ซึ่งจะต้องอยู่บนพื้นฐานของความต้องการและมุมมองของนักกีฬา โค้ช นักวิจัย และนักวิทยาศาสตร์ โดยประเด็นสำคัญที่ต้องพิจารณานั้นเกี่ยวกับบทบาทของการวิจัยในกีฬาสมัยใหม่และวิทยาศาสตร์การออกกำลังกาย และการวิจัยทางสรีรวิทยามีผลกับสภาพแวดล้อมและสมรรถนะของการเล่นกีฬาและการออกกำลังกายได้อย่างไร ซึ่งผู้วิจัยจำเป็นต้องพิจาณาประเด็นที่สำคัญเหล่านี้

การศึกษาและการวัดผลเพื่อการประเมินทางสรีรวิทยาของการออกกำลังกายเริ่มมีบทบาทและมีความสำคัญมากขึ้นในการพัฒนาของวงการวิทยาศาสตร์เพื่อการกีฬาและเพื่อสุขภาพ การประเมินผลของวิธีการวัดควรเป็นส่วนสำคัญของการฝึกพัฒนาศักยภาพของนักกีฬา โครงการสนับสนุนทางวิทยาศาสตร์ และควรดำเนินการอย่างเป็นระบบให้มากขึ้น นอกจากนี้การประเมินวิธีการวัดควรสะท้อนถึงกิจกรรมความเคลื่อนไหวและความต้องการอื่น ๆ ของกีฬาหรือนักออกกำลังกายด้วย ดังนั้นผู้วิจัยหรือผู้สนใจที่จะศึกษา ควรมีความเข้าใจอย่างละเอียดเกี่ยวกับกลไกการปลดปล่อยพลังงานโดยเฉพาะผลจากการทดสอบทางสรีรวิทยา (Bird and

Davison, 1997) เหตุผลการทดสอบทางสรีรวิทยาของการออกกำลังกาย ประกอบด้วย 1) เพื่อการประเมินเบื้องต้นในจุดแข็งและจุดอ่อนของผู้เข้าร่วมกิจกรรมทดสอบในบริบทของกีฬาหรือกิจกรรมต่างๆ ที่พวกเขามีส่วนร่วม ข้อมูลเหล่านี้สามารถเอาไว้ใช้ในการออกแบบและการใช้เป็นโปรแกรมการฝึกพัฒนาศักยภาพของนักกีฬาหรือผู้ออกกำลังกายได้ 2) เพื่อการประเมินประสิทธิผลของโปรแกรมการฝึกอบรมเพื่อดูว่าผลจากการปฏิบัติหรือการฟื้นฟูสมรรถภาพ ได้มีการปรับปรุงและมีการปรับตัวทางสรีรวิทยาตามที่ตั้งใจไว้หรือไม่ 3) เพื่อการประเมินภาวะสุขภาพของนักกีฬาหรือผู้ออกกำลังกาย ซึ่งนี้อาจเป็นส่วนหนึ่งของโครงการร่วมกับเจ้าหน้าที่ห้องทดสอบคลินิกได้ 4) เพื่อการให้ความช่วยเหลือเกี่ยวกับการวิเคราะห์การเคลื่อนไหว (Ergogenic) ไว้สำหรับการกำหนดเป้าหมายระยะสั้นสำหรับการปรับปรุงและการพัฒนาสมรรถนะของนักกีฬาและผู้ออกกำลังกาย 5) เพื่อช่วยในการคัดเลือกหรือการระบุความพร้อมในการฝึกอบรมหรือการแข่งขัน 6) เพื่อการพัฒนาความรู้ความเข้าใจเกี่ยวกับกีฬาหรือกิจกรรม ซึ่งจะเป็นประโยชน์ของโค้ช นักกีฬาได้ในอนาคตและรวมถึงนักวิทยาศาสตร์การกีฬาด้วย 7) เพื่อตอบคำถามการวิจัย ในด้านการแพทย์และการใช้ประโยชน์ในการทดสอบการออกกำลังกาย

เมื่อกล่าวถึงรูปแบบการวิจัยทางวิทยาศาสตร์การกีฬา สามารถแบ่งรูปแบบได้โดยทั่วไปเป็น 2 รูปแบบหลัก ประกอบด้วย การวิจัยเชิงปริมาณเป็นกระบวนการที่มีรูปแบบเฉพาะทางการมีวัตถุประสงค์และเป็นระบบ ซึ่งมุ่งใช้ข้อมูลตัวเลขเพื่อหาข้อมูลหรือผลลัพธ์ ที่เกี่ยวข้องเพื่อการทดสอบสมมุติฐานหรือเพื่อค้นหาความสัมพันธ์ โดยทั่วไปสามารถกล่าวได้ว่าเป็นประเภทของการวิจัยเชิงอนุมาน ซึ่งหมายความว่านักวิทยาศาสตร์การกีฬาจะเริ่มต้นจากสมมติฐานและเริ่มสังเกตการณ์เพื่อพิสูจน์สมมติฐาน โดยการตั้งสมมติฐานนั้นอาจถูกออกแบบเพื่อสร้างความแตกต่าง ความสัมพันธ์ หรือความเป็นเหตุเป็นผล อีกรูปแบบนึงคือ การวิจัยเชิงคุณภาพเป็นรูปแบบอัตนัยและเกี่ยวข้องกับคำมากกว่าตัวเลข เกี่ยวข้องกับความรู้สึก ความคิดเห็นและอารมณ์ และให้ความสำคัญกับการอธิบายว่าทำไมมากกว่าการค้นหาเชิงประมาณหรือจำนวน หมายความว่าสมมติฐานสามารถถูกพัฒนาผ่านการวิจัย ที่มีความพยายามที่จะอธิบายความแตกต่าง ความสัมพันธ์หรือความเป็นเหตุเป็นผล อย่างไรก็ตามข้อมูลเชิงคุณภาพยังสามารถสร้างเป็นข้อมูลเชิงปริมาณได้

โดยปัจจุบันการวิจัยทางวิทยาศาสตร์การกีฬามีการพัฒนาและก้าวหน้าอย่างต่อเนื่อง มีการนำองค์ความรู้ที่ได้รับจากการวิจัยเหล่านี้มาประยุกต์ใช้เพื่อในการดูแลและพัฒนาศักยภาพของนักกีฬาและผู้ออกกำลังกายกันมากขึ้น อีกทั้งมีผู้สนใจและนักวิจัยรุ่นใหม่ที่ต้องการศึกษาและวิจัยเพื่อใช้ในการพัฒนาองค์ความรู้ด้านวิทยาศาสตร์การกีฬาจำนวนที่เพิ่มขึ้น องค์ความรู้ที่ได้รับจากการวิจัยจะมีความน่าเชื่อถือและถูกต้องตามหลักกระบวนการทางวิทยาศาสตร์ควรต้องได้รับการพิจารณาประเด็นสำคัญจากผู้วิจัยหรือผู้สนใจที่จะศึกษาทั่วไป งานวิจัยที่มีคุณภาพนั้นผู้วิจัยหรือผู้ที่ศึกษานั้นจำเป็นต้องดำเนินการวิจัยตามกระบวนการอย่างถูกต้อง ประกอบด้วย ความเข้าใจในหลักการและระเบียบวิธีการวิจัยทางวิทยาศาสตร์การกีฬา ความเข้าใจและสามารถเลือกใช้วิธีการทางสถิติเบื้องต้นในการวิเคราะห์ข้อมูลได้ ความเข้าใจและความสามารถในการออกแบบการวิจัย โดยสามารถเลือกใช้เครื่องมือที่เหมาะสมกับปัญหาวิจัยได้อย่างเหมาะสมและได้มาตรฐานทางวิชาชีพด้านวิทยาศาสตร์การกีฬา

อย่างไรก็ตาม งานวิจัยที่มีคุณภาพจำเป็นต้องพิจาณาอย่างละเอียดในประเด็นหรือปัจจัยที่เกี่ยวข้องที่อาจสามารถส่งผลต่อคุณภาพของงานวิจัยได้ โดยเฉพาะอย่างยิ่ง ในการวิจัยทางด้านสรีรวิทยาของการออกกำลังกายนั้น ผลการวิจัยที่จะเกิดขึ้นอย่างมีคุณภาพและสามารถเป็นที่ยอมรับจากวงการวิทยาศาสตร์การกีฬาได้จำเป็นต้องพิจารณาในส่วนของ ความถูกต้อง ความน่าเชื่อถือ ความแม่นยำ และความเที่ยงตรง ดังนั้นสำหรับบทความนี้ มุ่งเน้นที่จะศึกษาในประเด็น ความถูกต้อง ความน่าเชื่อถือ ความแม่นยำ และความเที่ยงตรง ของการวิจัยทางด้านสรีรวิทยาของการออกกำลังกาย เพื่อเป็นประโยชน์สำหรับนักวิจัยและผู้สนใจทางด้านวิทยาศาสตร์การกีฬาไว้สำหรับการพิจารณาถึงปัจจัยที่มีอิทธิพลต่อคุณภาพของงานวิจัยทางด้านสรีรวิทยาของการออกกำลังกายได้

แนวคิดของ ความน่าเชื่อถือ ความถูกต้อง ความแม่นยำ ความเที่ยงตรง

งานวิจัยที่มีคุณภาพนั้น ควรพิจารณาถึงปัจจัยที่มีอิทธิพลต่อคุณภาพของงานวิจัยทางวิทยาศาสตร์การกีฬา โดยเฉพาะคุณภาพของงานวิจัยทางด้านสรีรวิทยาของการออกกำลังกายนั้น มีความเกี่ยวข้องกับ ความถูกต้อง ความน่าเชื่อถือ ความแม่นยำ และความเที่ยงตรง

ความถูกต้อง (Validity)

ความถูกต้อง (Validity) เป็นสิ่งสำคัญในการวิจัยเพราะการแสดงออกมาให้เห็นว่าค่าหรือผลลัพธ์ที่ได้จากการวัดเป็นค่าจริง ตรงตามสิ่งที่ผู้วิจัยวางแผนไว้ก่อนที่จะวัด ประเภทของความถูกต้องมีอยู่ 2 ประเภทคือ ความถูกต้องภายในและความถูกต้องภายนอก ความถูกต้องภายในเกี่ยวกับผลของการศึกษาสามารถนำมาประกอบกับการศึกษา การทดลอง หรือการทดสอบที่แตกต่างกัน นั่นหมายความว่าสำหรับงานวิจัยนั้นสามารถอ้างได้ในความถูกต้องภายในของงานวิจัยที่ศึกษา แต่ต้องได้รับการตรวจสอบเพื่อยืนยันให้แน่ใจว่าการวิจัยนี้ได้มีการควบคุมปัจจัยทุกอย่างที่อาจส่งผลต่อผลการศึกษา ความถูกต้องภายนอกเกี่ยวข้องกับผลการวิจัยนั้นๆ สามารถนำไปประยุกต์ใช้กับโลกแห่งความเป็นจริงได้หรือไม่

ความถูกต้อง คือขอบเขตของการวัดหรือการทดสอบและการวัดเกี่ยวข้องกับอะไรบ้างหรืออะไรที่หมายถึงการวัด ตัวอย่างเช่นการประเมินกลไก ที่อธิบายถึงความทนทานหรือความแข็งแรงและการใช้โครงสร้างทางกายภาพที่เหมาะสม ดังนั้นการวัดเกี่ยวข้องกับ ผลลัพธ์ของระดับความเข้มข้นสูงสุดหรือปริมาณการออกกำลังกายทั้งหมด

ความน่าเชื่อถือ (Reliability)

ความน่าเชื่อถือ (Reliability) เกี่ยวข้องกับว่า ถ้าหากนักวิจัยทำวิจัยซ้ำอีกครั้ง จะได้ผลลัพธ์ที่เหมือนหรือคล้ายคลึงกันหรือไม่ อย่างไรก็ตามความน่าเชื่อถือสามารถอ้างโดยไม่จำเป็นต้องมีผลลัพธ์ที่ถูกต้อง ตัวอย่างเช่นหากนักวิจัยถามคำถามผิดในการวิจัยบ่อยๆผลจากการวิจัยนั้นก็จะได้รับคำตอบที่ผิดเสมอเช่นกัน ซึ่งหมายความว่าการทดสอบนั้นมีความน่าเชื่อถือได้เพราะนักวิจัยได้รับคำตอบผิดเดียวกันแม้ว่าจะไม่ใช่คำตอบที่ถูกก็ตาม ในการวิจัยเชิงปริมาณความน่าเชื่อถือ อาจบ่งบอกได้จากผู้วิจัยคนหนึ่งที่ทำการทดสอบเดียวกันโดยบุคคลเดียวกัน ในหลายโอกาสและได้ผลเหมือนกันหรือคล้ายกัน หรืออาจเป็นนักวิจัยที่แตกต่างกันที่ทำแบบทดสอบ

เดียวกันและได้ผลเหมือนกันหรือคล้ายคลึงกัน ในการวิจัยเชิงคุณภาพความน่าเชื่อถือเกี่ยวข้องกับนักวิจัยคนเดียวกันที่สามารถจัดวางและเรียบเรียงข้อมูลที่ได้จากการวิจัยจากกระบวนการวิเคราะห์ข้อมูลและแปลผลจากสิ่งที่ศึกษาและค้นพบ ไว้ในหมวดเดียวกันในโอกาสที่ต่างกันหรือนักวิจัยที่ต่างกัน หรือการวางผลลัพธ์จากการวิจัยไว้ในหมวดเดียวกันหรือคล้ายคลึงกันโดยมีปัจจัยบางอย่างที่ควรคำนึงถึงซึ่งอาจส่งผลต่อความน่าเชื่อถือของงานวิจัย อาทิ ความคลาดเคลื่อนที่อาจเกิดขึ้นได้เมื่อนักวิจัยไม่ทราบวิธีการใช้อุปกรณ์หรือเครื่องมือการวัดได้อย่างถูกต้อง หรืออุปกรณ์หรือเครื่องมือการวัดอาจได้รับการบำรุงรักษาที่ไม่ดี และอุปกรณ์หรือเครื่องมือการวัดที่ไม่ถูกต้องซึ่งอาจถูกนำมาใช้ ส่วนใหญ่ความผิดพลาดจากการใช้เครื่องมือ ทำให้เกิดความไม่น่าเชื่อถือขึ้นนั้น เป็นเหตุมาจากตัวบุคคล(หรือผู้ทำการวัด) เช่น การอ่านค่าผิด การปรับแต่งหรือใช้เครื่องมือผิดหรือการคำนวณที่ผิด ล้วนเป็นผลมาจากความผิดพลาดจากบุคคล (Human Errors)

ความน่าเชื่อถือสามารถแบ่งได้สองประเภทคือความน่าเชื่อถือได้ระหว่างผู้วิจัย (Inter-Researcher Reliability) และความน่าเชื่อถือได้ในการวัดซ้ำ (Test-Retest Reliability) ความน่าเชื่อถือระหว่างผู้วิจัยพบว่าผู้วิจัยที่ต่างกันในสถานการณ์เดียวกันจะได้รับผลได้ผลเหมือนกันหรือใกล้เคียงกัน ตัวอย่างของความน่าเชื่อถือระหว่างผู้วิจัยคือ ประเด็นด้านการประเมินส่วนประกอบของร่างกาย เมื่อผู้วิจัยมีการเรียนรู้ที่จะใช้เทคนิค Calipers skinfold (หรือการวัดปริมาณไขมันด้วยเครื่องหนีบ) ในการประเมินส่วนประกอบของร่างกาย ผู้วิจัยจะสามารถวัดได้อย่างแม่นยำจากองค์ประกอบหรือตัวบุคคลที่มีความถูกต้อง ถ้าหากผลลัพธ์จากผู้วิจัยมีค่าต่างกันจะไม่สามารถอ้างว่าได้รับความน่าเชื่อถือระหว่างนักวิจัย

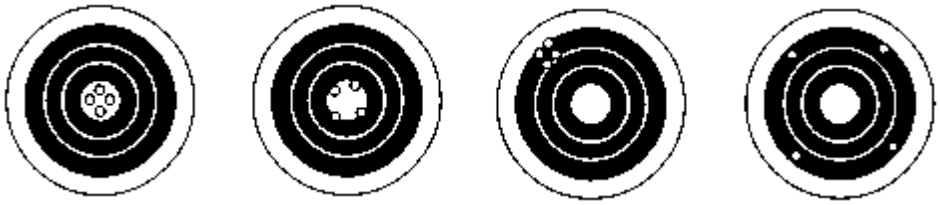
ความน่าเชื่อถือเกี่ยวข้องกับการทดสอบความถูกต้องในการวัดซ้ำๆกันเพื่อใช้ในการสอบเทียบเกี่ยวข้องกับการทำแบบทดสอบเดียวกันในแต่ละโอกาสและได้ผลเหมือนกันหรือใกล้เคียงกัน ตัวอย่างของปัญหาความน่าเชื่อถือในการทดสอบซ้ำที่พบในการวิจัยของการกีฬาหรือการออกกำลังกายคือ การวัดอัตราการเต้นของหัวใจ ซึ่งอัตราการเต้นของหัวใจอาจได้รับผลกระทบจากปัจจัยต่างๆเช่นอุณหภูมิ ช่วงเวลาของวัน อาหาร รูปแบบหรือประสิทธิภาพของการนอนหลับ ระดับการออกกำลังกาย และแอลกอฮอล์ ถ้าวัดอัตราการเต้นของหัวใจกับคนๆเดียวกันและในเวลาเดียวกัน แต่ในวันต่างๆอาจได้รับค่าจากการวัดที่ต่างกันได้

ความแม่นยำ (Accuracy)

ความแม่นยำ (Accuracy) เกี่ยวข้องกับการวัดได้ใกล้เคียงกับ ค่าการตรวจมาตรฐานสูงสุด (Gold Standard) หรือค่าที่ตั้งใจจะวัดหรือค่าจริงหรือไม่ อีกนัยหนึ่งกล่าวได้ว่า ความแม่นยำ คือความถูกต้องของค่าที่วัดได้ เป็นความใกล้เคียงกับค่าจริงหรือใกล้เคียงกับค่าจากเครื่องมือมาตรฐาน ยกตัวอย่างเช่น หากมีการวัดน้ำหนักของนักมวยก่อนขึ้นชกจริง ซึ่งสามารถวัดค่าน้ำหนักได้ 70.1 kg เมื่อเทียบกับน้ำหนักจริงของนักมวย ซึ่งมีค่าน้ำหนักที่ 70 kg สามารถกล่าวได้ว่าค่าที่ได้จากการวัดจากเครื่องชั่งน้ำหนักของฝ่ายการจัดการแข่งขัน มีความแม่นยำหรือความถูกต้อง

ความเที่ยงตรง (Precision)

ความเที่ยงตรง (Precision) หมายถึง ความใกล้เคียงกันในกลุ่มของค่าที่ทำการวัด เมื่อเกี่ยวข้องกับการดำเนินการในการตั้งค่าการวิจัย ในกระบวนการวัดนั้นจะมีบางส่วนที่คาดเดาไม่ได้ มีระดับของความไม่แน่นอนเกิดขึ้นและเกี่ยวข้องกับขนาดของความเที่ยงตรงของเครื่องมือวัด ความเที่ยงตรงเกี่ยวข้องกับการปรับแต่งของกระบวนการวัด หรือที่เกี่ยวข้องกับผลของความแตกต่างที่เกิดจากเครื่องมือวัดที่สามารถตรวจจับได้เล็กน้อย และความเที่ยงตรงเกี่ยวข้องกับกระบวนการทำซ้ำ สามารถเกี่ยวกับการวัดผลให้มีความแม่นยำและเที่ยงตรง ให้คิดถึงกีฬายิงธนู แล้วนักธนูต้องยิงให้เข้าเป้าบนกระดานยิงธนู ทำความเข้าใจในภาพที่ 1 แสดงถึงจำนวนลูกศรที่ถูกยิงเพื่อให้เข้าเป้า ทั้งหมดนั้นจะถูกแสดงอยู่บนกระดานยิงธนู ซึ่งแสดงถึงความเที่ยงและแม่น อย่างไรก็ตามหากลูกศรนั้นถูกยิงกระจัดกระจายบนกระดานตามทิศทางต่างๆ สามารถกล่าวได้ว่าการยิงธนูนั้นไม่มีความเที่ยงและไม่มีความแม่น แสดงในแผนภาพด้านล่าง



ภาพที่ 1: การอธิบายถึงความแม่นและความเที่ยงของลูกศรที่อยู่บนเป้ายิงธนู
ที่มา:ดัดแปลงจาก http://explorable.com

เกณฑ์ในการทดสอบของการวิจัยที่เกี่ยวกับสรีรวิทยาการออกกำลังกาย (TEST CRITERIA)

เป็นที่ยอมรับกันว่า การมีประสิทธิภาพของการกระบวนการวัดผลหรือการประเมินวิธีการวัดหรือการทดสอบควรมีลักษณะดังนี้ ความจำเพาะ และมีความถูกต้องอย่างมีเหตุผล และกระบวนการวัดผลหรือการประเมินการทดสอบที่เกิดขึ้นควรมีความไวต่อการเปลี่ยนแปลงที่เกิดจากสมรรถนะของผู้ถูกทดสอบ กระบวนการวัดผลเป็นการหาค่าทั้งเชิงปริมาณและเชิงคุณภาพ ของสิ่งที่วัดหรือสภาพแวดล้อมที่กำลังพิจารณา เครื่องมือวัดที่ใช้ถูกใช้เพื่อหาค่าหรือขนาดทางปริมาณ หรือความเปลี่ยนแปลงต่าง ๆ ของสิ่งที่กำลังพิจารณา ในส่วนของความคลาดเคลื่อน อาจอธิบายได้ว่าเป็นผลจากความเบี่ยงเบนของค่าที่วัดได้จากค่าจริง

กระบวนการวัดควรมีความจำเพาะ (Specificity)

การประเมินการทดสอบควรจำลองรูปแบบของการออกกำลังกายที่อยู่ภายใต้การตรวจสอบข้อเท็จจริง นี่เป็นความท้าทายอย่างหนึ่งในการวิจัยแขนงนี้ อาทิ ในกิจกรรมที่เกี่ยวข้องกับความเร็วหรือการวิ่งด้วยความเร็ว หรือกิจกรรมหลายอย่างเช่นการแข่งขันในสนามกีฬาและกีฬาแร็กเก็ต ที่มีการเปลี่ยนแปลงความเร็วและทิศทางการเคลื่อนที่ ปัจจัยที่ควรพิจารณาในการออกแบบวิธีการหรือรูปแบบของการทดสอบ (Atkinson and Nevill, 1998) ได้แก่

1) กลุ่มกล้ามเนื้อที่เกี่ยวข้องกับประเภทของแต่ละกิจกรรมและช่วงของการเคลื่อนไหวที่เกี่ยวข้องซึ่งจำเป็นต้องพิจารณา

2) ความเข้มข้นและระยะเวลาของกิจกรรมการออกกำลังกาย

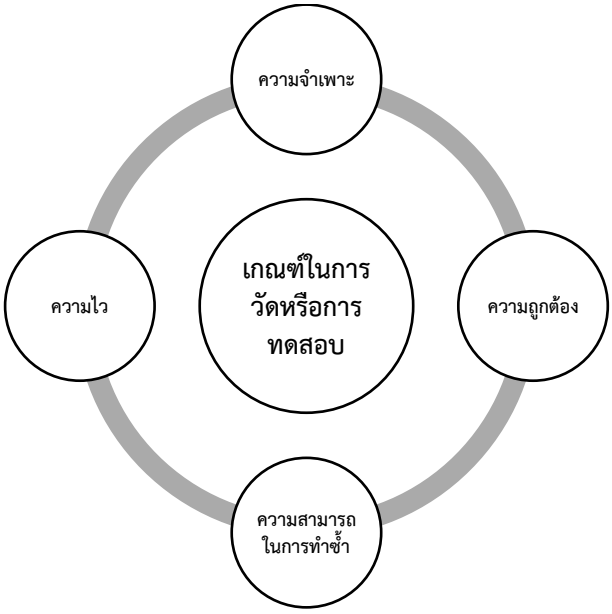
3) ระบบการใช้พลังงาน

4) แรงต้านที่อาจจะเกิดขึ้น

ดังนั้นควรใช้เครื่องวัดการทรงตัวแบบเฉพาะกิจกรรมและอาจต้องมีการออกแบบเพื่อตอบสนองต่อความต้องการนั้นๆ ในทำนองเดียวกันการดำเนินวิจัยภาคสนาม เมื่อเทียบกับวิธีปฏิบัติในห้องปฏิบัติการ อาจปรับปรุงลักษณะของรูปแบบการเคลื่อนไหว ดังนั้นไม่จำเป็นว่าต้องเหมือนกับรูปแบบของการทดลองที่เคยทำมาก่อนหน้านี้ แต่ควรพิจารณาโดยการทดสอบ ตัวอย่างเช่นในการทดสอบเชิงการวินิจฉัยอาจมีการเลือกที่มีการพิจารณาอย่างรอบคอบในการเลือกรูปแบบการออกกำลังกายที่ไม่จำเป็นต้องสะท้อนถึงกิจกรรมในประจำวัน อาทิ เครื่องวัดระยะของการทรงตัวของวงจรอาจจัดอยู่ในประเภทนี้ แต่การใช้งานจะทำให้ได้รับสัญญาณที่ดีขึ้นในการวัดคลื่นไฟฟ้าหัวใจและระบบการเผาผลาญ

กระบวนการวัดควรมีความถูกต้อง (Validity)

กระบวนการวัดควรมีความถูกต้อง (Validity) หมายความว่า ขอบเขตที่การทดสอบจำเป็นต้องวัดอะไรบ้างถึงจะสามารถอธิบายถึงวิธีการวัดนั้นได้ หรืออะไรที่หมายถึงการวัด ตัวอย่างเช่นการประเมินกลไกการทดสอบที่สามารถอธิบายถึงความอดทนและการใช้โครงสร้างทางกายที่เหมาะสม (Atkinson et al., 2005a) ซึ่งผลลัพธ์ที่ได้จากการทดสอบ คือระดับความเข้มข้นขอการออกกำลังกายสูงสุดหรือปริมาณการออกกำลังกายทั้งหมด



ภาพที่ 2: เกณฑ์ที่ใช้เพื่อการพิจารณาในการวัดหรือการทดสอบในงานวิจัยทางสรีรวิทยาของการออกกำลังกาย

กระบวนการวัดควรมีความสามารถในการทำซ้ำได้ (Reproducibility)

กระบวนการวัดหรือการทดสอบควรมีความสามารถในการทำซ้ำได้ (Reproducibility) ข้อกำหนดที่สำคัญสำหรับข้อมูลที่ได้จากการวัดหรือการทดสอบก็คือมีความสามารถในการทำซ้ำได้ การอภิปรายที่ผลการวิจัยยังคงดำเนินต่อ เกี่ยวกับเมตริกหรือหน่วยสำหรับการวัดปริมาณต่างๆ ที่ประเมินความสามารถในการทำซ้ำได้อย่างเหมาะสมที่สุด (Atkinson and Nevill, 1998) ดังนั้นนักวิทยาศาสตร์การกีฬาของการออกกำลังกายต้องมีความเข้าใจและเชื่อมั่นอย่างมากต่อกระบวนการวัดเหล่านี้และรับรู้ถึงข้อดีและข้อเสียของแต่ละวิธีการวัด

ในสาระสำคัญความแปรปรวนที่เกิดจากกระบวนการวัด สามารถนำมาประกอบกับแหล่งทางเทคนิคและทางชีวภาพ ความแม่นยำและความเที่ยงตรงของเครื่องมือวัด ประกอบกับทักษะของผู้วิเคราะห์หรือผู้วิจัย จึงจำเป็นต้องให้ความสำคัญกับขั้นตอนในการสอบเทียบอย่างยิ่ง ประกอบกับความแปรปรวนทางชีวภาพที่มีความไม่แน่นอนและที่เป็นรูปแบบวัฏจักร รวมถึงความรู้และความเข้าใจเกี่ยวกับระดับความคลาดเคลื่อนมีบทบาทสำคัญที่ต้องทำความเข้าใจและสามารถอธิบายได้ เป็นผลให้ตัวบ่งชี้ของความคลาดเคลื่อนในการทดสอบนั้นถูกนำไปเป็นข้อกำหนดได้ หากข้อมูลจากที่ได้จากตัวบ่งชี้ของความคลาดเคลื่อนสามารถให้คำอธิบายได้มากขึ้น ซึ่งรวมถึงการเปรียบเทียบผลการทดสอบกับมาตรฐานหรือผลการทดสอบของนักวิจัยคนอื่น ๆ

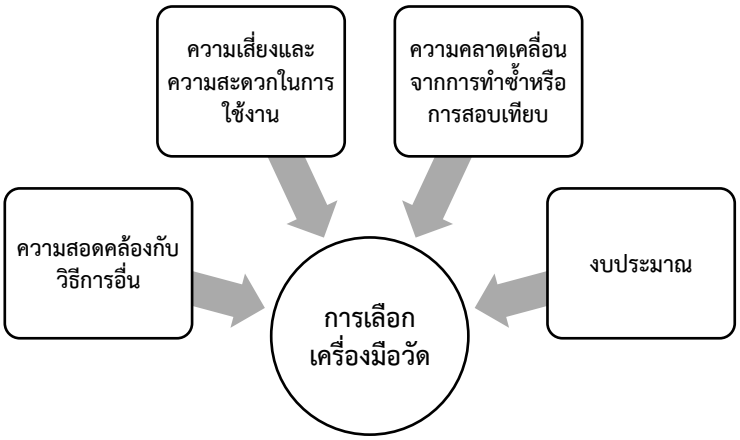
กระบวนการวัดควรมีความไว (Sensitivity)

ความไว (Sensitivity) คือ ขอบเขตของการวัดทางสรีรวิทยาแสดงถึงการปรับปรุงสมรรถนะเห็นได้ชัดว่าความสามารถในการทำซ้ำเกี่ยวข้อง แต่ความไวอาจเป็นหัวใจสำคัญของกระบวนการวัดหรือการทดสอบ ความไวเป็นตัวชี้วัดที่สำคัญในการทำความเข้าใจเกี่ยวกับกลไกของเครื่องมือวัด ความเที่ยงตรงและความแม่นยำของเครื่องมือวัดก็เพื่อที่จะสะท้อนถึงกลไกเหล่านี้ เป็นไปได้ว่าการประเมินลักษณะทางสรีรวิทยาของนักกีฬาและนักออกกำลังกายจะยังคงเป็นส่วนสำคัญของโปรแกรมทางวิทยาศาสตร์การกีฬา ในทำนองเดียวกันอาจเป็นไปได้ว่าวิธีการประเมินจะยังคงต้องได้รับการพัฒนาและการปรับปรุงให้ดีขึ้นเนื่องจากฐานองค์ความรู้มีการเติบโตขึ้น องค์ความรู้เหล่านี้จะช่วยเพิ่มระดับความไวที่มีต่อการวัดทางสรีรวิทยา ซึ่งการวัดนี้สามารถอธิบายการเปลี่ยนแปลงในสมรรถนะของนักกีฬาและผู้ออกกำลังกายได้ (Atkinson and Nevill, 1998) ดังนั้นในการประเมินกระบวนการวัดหรือการทดสอบควรมีเหตุผล ซึ่งจะเป็นประโยชน์สำหรับนักกีฬาที่ต้องการให้เกิดความรู้ความชำนาญ และเกิดประสบการณ์ในการออกกำลังกายที่เพิ่มขึ้น

ความสอดคล้องในวิธีการและความคลาดเคลื่อนจากการวัด ในทางสรีรวิทยาของการออกกำลังกาย (Method Agreement and Measurement Error in the Physiology of Exercise)

Atkinson et al. (2005a) ได้ให้ข้อเสนอแนะกับ นักสรีรวิทยาด้านการออกกำลังกายว่าจำเป็นต้องเลือกเครื่องมือวัดหรือเครื่องมือทดสอบที่เหมาะสมที่สุด ก่อนที่จะเริ่มรวบรวมข้อมูลจากนักกีฬาหรือผู้เข้าร่วมในการวิจัยหรือการทดสอบนั้น ๆ หลักเกณฑ์สำคัญที่ใช้ประกอบกับการเลือกเครื่องมือวัดหรือเครื่องมือทดสอบประกอบด้วย ดังนี้:

1. ระดับของความเสี่ยงจากการใช้งานและระดับความสะดวกในการใช้งานที่เหมาะสม
2. งบประมาณที่มีอยู่
3. ระดับความคลาดเคลื่อนจากการทำซ้ำ
4. ระดับความสอดคล้องในวิธีการที่แตกต่างกันกล่าวคือ การตรวจว่าวิธีการที่แตกต่างกันนั้น จะได้มาซึ่งข้อมูลที่มีความสัมพันธ์กันหรือมีความสอดคล้องกันหรือไม่ จำเป็นต้องให้ใช้วิธีการอื่นหรือวิธีการที่แตกต่างกันไป ซึ่งอาจหมายถึง วิธีการที่ซึ่งอาจเป็นอันตรายมากขึ้น ไม่สะดวก หรือแพงกว่า



ภาพที่ 3: หลักเกณฑ์สำคัญที่ควรคำนึงถึงในการเลือกใช้เครื่องมือวัด

จากภาพที่ 3 แสดงถึงหลักเกณฑ์สำคัญที่ผู้วิจัยควรคำนึงถึงในการเลือกใช้เครื่องมือวัด อย่างไรก็ตามก็ยังไม่เพียงพอต่อการตัดสินใจเลือกที่เหมาะสม ผู้วิจัยควรพิจารณาเพิ่มเติมในประเด็นที่เกี่ยวข้องกับการวัดลักษณะทางสรีรวิทยาของมนุษย์บางส่วนที่ยังมีข้อผิดพลาดในกระบวนการทำซ้ำหรือการทวนสอบผลการวัดเนื่องมาจากการเปลี่ยนแปลงทางชีวภาพตามธรรมชาติ อาทิ ปัจจัยบุคคลหรือสภาพแวดล้อมในระหว่างการวัดหรือการทดสอบ ในการประเมินเครื่องมือวัดซึ่งรวมทั้ง ที่ผู้วิจัยกำลังทำการศึกษาโดยการออกแบบการทดสอบด้วยผู้วิจัยเองหรือจากการที่ผู้วิจัยได้จากการทบทวนวรรณกรรม จากหนังสือที่เกี่ยวข้องซึ่งผ่านการตีพิมพ์ในวารสารทางวิทยาศาสตร์ โดยขั้นตอนการเลือกเครื่องมือวัดประกอบด้วย การตรวจสอบความสอดคล้องระหว่างวิธีการวัด หรือการศึกษาเพื่อเปรียบเทียบวิธีการ ความคลาดเคลื่อนในกระบวนการทำซ้ำกล่าวคือการทำซ้ำหรือการสอบเทียบผลการวัดซึ่งกล่าวได้ว่าเป็นการทดสอบความน่าเชื่อถือของข้อมูลที่ได้จากการวัด การตรวจสอบความคลาดเคลื่อนเชิงระบบ การตรวจสอบความคลาดเคลื่อนแบบไม่แน่นอน และความแม่นยำทางสถิติ

เป็นสิ่งสำคัญที่จะต้องเข้าใจว่า เครื่องมือวัดที่มีราคาแพงที่สุดและเครื่องมือวัดที่ต้องใช้วิธีการสอดเข้าไปในร่างกาย อาจไม่ได้แสดงว่า ค่าความคลาดเคลื่อนจากการวัดหรือการทดสอบซ้ำจะมีค่าน้อยที่สุด นอกจากนี้วิธีการวัดทั้งหมดที่ใช้ในการวัดลักษณะทางสรีรวิทยาของมนุษย์ยังมีความคลาดเคลื่อนบางส่วนที่พบ ซึ่งอาจเกิดได้จากกระบวนการทำซ้ำเนื่องมาจากการเปลี่ยนแปลงทางชีวภาพตามธรรมชาติ ตัวอย่างเช่น การใช้วิธีที่เรียกกันว่า “การตรวจสอบตามมาตรฐานสูงสุดของดักลาส” เกี่ยวข้องกับความคลาดเคลื่อนในกระบวนการทำซ้ำ ที่เกี่ยวเนื่องมาจากความแปรปรวนของจลนพลศาสตร์ของการบริโภคออกซิเจนเข้าไปในร่างกายของมนุษย์ในระหว่าง

การออกกำลังกาย (Atkinson et al., 2005a) กล่าวคือ เป็นการตรวจวินิจฉัยในรูปแบบคลินิก ที่ดีที่สุดภายใต้เงื่อนไขที่เหมาะสม ทั้งนี้อาจไม่ใช่การตรวจวินิจฉัยที่ได้ความแม่นยำที่สุด

ในทำนองเดียวกันในขณะที่การเปรียบเทียบเครื่องวัดความดันโลหิตอัตโนมัติแบบใหม่กับเครื่องบันทึกชีพจร วิธีการนี้เกี่ยวข้องกับความคลาดเคลื่อนในการทำซ้ำ (Bland and Altman, 1999) การพิจารณาถึงการขยายหรืออัตราที่เพิ่มขึ้นของความแปรปรวนทางชีวภาพตามธรรมชาตินี้มีผลต่อการวิเคราะห์ในสมรรถนะของเครื่องมือวัดทางสรีรวิทยาด้วยกระบวนการตรวจสอบความคลาดเคลื่อนในการทำซ้ำควรมีอยู่ในการตรวจสอบความสอดคล้องระหว่างเครื่องมือวัดต่างๆ และควรได้รับการประเมินสำหรับค่าความคลาดเคลื่อนในการทำซ้ำ หากผ่านการตรวจสอบและการวิเคราะห์ดังกล่าวแล้ว นักวิจัยก็สามารถตรวจสอบได้เช่นกัน นักวิจัยสามารถสร้างข้อสรุปที่ชัดเจนได้ แม้กระทั่งแหล่งที่มาของความไม่สอดคล้องกันระหว่างวิธีการวัดที่มีความแตกต่างกัน (Bland and Altman, 2003; Atkinson et al., 2005a)

ในบางแง่มุมของ “การวิเคราะห์ความถดถอยโดยวิธีกำลังสองน้อยที่สุด (LSR)” เป็นวิธีการประมาณค่า a และ b จากข้อมูล (x1,y1), (x2,y2),....(xn, yn) เพื่อให้ได้เส้นตรงที่เข้ากับข้อมูลได้ดีที่สุด วิธีกำลังสองน้อยที่สุดจะได้ค่าประมาณ a และ b ที่ทำให้ความแตกต่างของค่าตัวแปรตามกับค่าที่คาดคะเนที่ได้จากสมการถดถอยมีค่าน้อยที่สุด ค่าสถิติ LSR ควรถูกใช้โดยมีการตรวจสอบในความสอดคล้องอย่างระมัดระวัง ระหว่างวิธีการวัดที่เกี่ยวข้องกับสรีรวิทยาการออกกำลังกาย มีความเป็นไปได้สูงว่าวิธีการวัดทางสรีรวิทยาทั้งสองนี้แสดงให้เห็นว่ามีความคลาดเคลื่อนในการทำซ้ำ เนื่องจากส่วนหลักของความผิดพลาดนี้เกิดจากการขยายหรืออัตราความแปรปรวนที่เพิ่มขึ้นจากแหล่งกำเนิดทางชีวภาพที่เกิดจากธรรมชาติ กล่าวได้ว่าความผิดพลาดนี้อาจเกิดขึ้นเมื่อมีการใช้วิธีการวัดอย่างใดอย่างหนึ่ง ซึ่งหมายความว่าสมมติฐานที่สำคัญสำหรับค่า LSR ซึ่งนำไปสู่การประมาณค่าทางสถิติที่มีความลำเอียง ในการประมาณค่าความชัน LSR และประมาณการค่าจุดตัด (Bland and Altman, 2003; Atkinson et al., 2005a) โดยค่าสถิติเหล่านี้ถูกนำไปใช้ในการอนุมานเกี่ยวกับความคลาดเคลื่อนเชิงระบบระหว่างวิธีการ แต่ค่าความชันและการตัดของเส้น ALSR จะไม่เป็นกลางหากวิธีการ 'Predictor' เกี่ยวข้องกับระดับการทำซ้ำที่ต่ำกว่าวิธีการอื่นหรือต่ำกว่าในความเป็นจริง ‘ตัวแปรคงที่’ นอกจากนี้ปรัชญาของการประมาณการของการวิเคราะห์ความถดถอยหรือการวิเคราะห์ความสัมพันธ์เชิงเส้น ไม่ได้สอดคล้องกับข้อเท็จจริงที่ว่านักวิจัยส่วนใหญ่ต้องการเลือก เครื่องมือวัดที่ดีที่สุดเป็นลำดับแรกเพื่อใช้ในการตรวจสอบ แทนที่จะใช้อีกวิธีการวัดเป็นส่วนหนึ่งของการศึกษาเพื่อการตรวจสอบความคลาดเคลื่อนเชิงระบบ

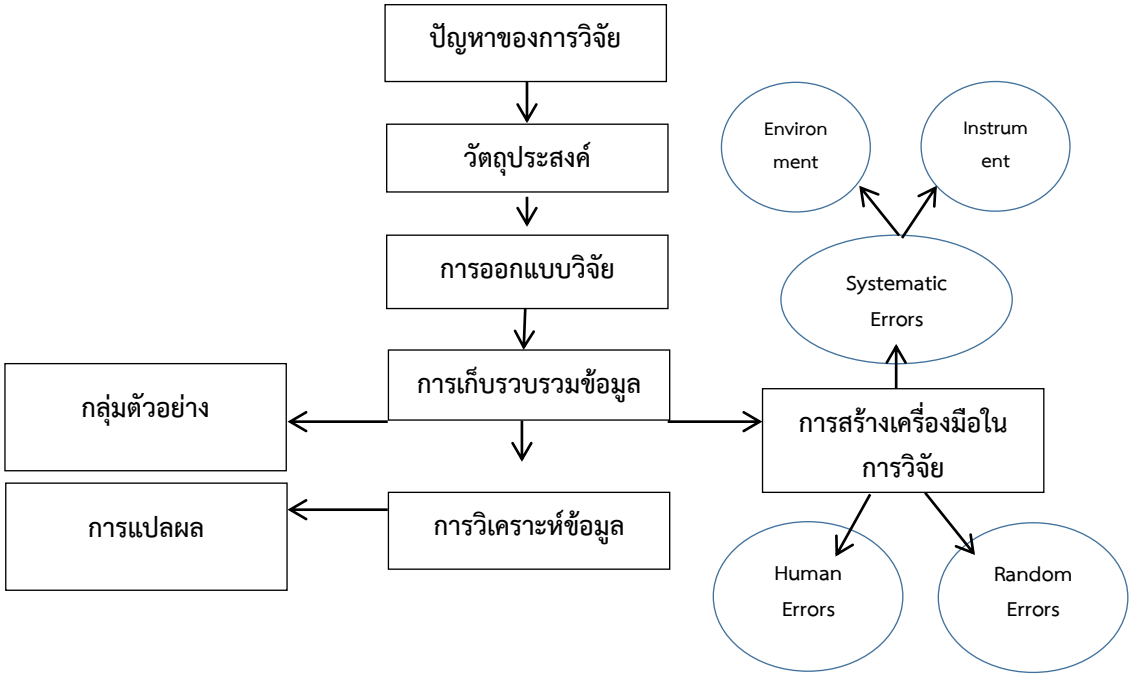
เช่นเดียวกับค่าสถิติอื่น ๆ ทั้งหมดที่ใช้ในการอธิบายความคลาดเคลื่อนและความสอดคล้องในกลุ่มประชากรที่มีความจำเพาะ ซึ่งกลุ่มประชากรที่ต่างกันอาจแสดงถึงระดับของความคลาดเคลื่อนที่แตกต่างกัน เนื่องจากแหล่งที่มาทางชีววิทยาตามธรรมชาติ บุคคลที่มีความแตกต่างกันเป็นตัวอย่างจากกลุ่มประชากรเดียวกันนี่อาจแสดงถึงระดับที่แตกต่างกันของความคลาดเคลื่อน ตัวอย่างเช่น บุคคลทั่วไปที่ถูกบันทึกค่าสูงที่สุดในทางสรีรวิทยา อาจแสดงถึงระดับของความคลาดเคลื่อนในการวัดที่มากด้วยเช่นกัน

ดังนั้นความคลาดเคลื่อนอาจทำให้เกิดความแตกต่างกันไปสำหรับแต่ละบุคคลในกลุ่มประชากร และความแม่นยำทางสถิติของการประมาณการความคลาดเคลื่อนของตัวอย่างนั้นเพียงพอหรือไม่นั้น ควรเป็นสิ่งที่ต้องพิจารณาอย่างละเอียด มีประเด็นเชิงปรัชญาอื่น ๆ ซึ่งสนับสนุนให้มีการใช้เทคนิคทางสถิติ ที่สามารถเอาไว้ใช้

ในการประเมินเครื่องมือวัดทางสรีรวิทยา (Atkinson and Nevill, 1998; Atkinson et al., 2005) อีกวิธีหนึ่งคือผู้วิจัยสามารถให้ความสำคัญในประเด็นที่สำคัญที่สุดของการศึกษาและการประเมินของวิธีการวัดผลหรือการทดสอบได้ โดยผู้วิจัย ผู้ศึกษาทั่วไปที่สนใจ สามารถประยุกต์ใช้รูปแบบของรายการตรวจสอบสำหรับนักสรีรวิทยาการออกกำลังกายนี้ได้ มีรายละเอียดดังนี้

แนวทางการตรวจสอบที่เกี่ยวกับ "ความสอดคล้องในวิธีการและความคลาดเคลื่อนจากการวัด" (A METHOD AGREEMENT AND MEASUREMENT ERROR CHECKLIST)

ผู้เขียนนำเสนอรายการตรวจสอบซึ่งอาจเป็นประโยชน์สำหรับนักสรีรวิทยาที่สนใจในการประเมินเครื่องมือวัดหรือเครื่องมือทดสอบซึ่งรวมทั้ง ที่ผู้วิจัยกำลังทำการศึกษาโดยการออกแบบการทดสอบด้วยผู้วิจัยเองหรือจากการที่ผู้วิจัยได้จากการทบทวนวรรณกรรม โดยขั้นตอนการเลือกเครื่องมือวัดหรือเครื่องมือการทดสอบนั้น ประกอบด้วย การตรวจสอบความสอดคล้องระหว่างวิธีการวัดหรือการศึกษาเพื่อเปรียบเทียบวิธีการความคลาดเคลื่อนในการทำซ้ำ การตรวจสอบความคลาดเคลื่อนเชิงระบบ การตรวจสอบความคลาดเคลื่อนแบบไม่แน่นอน (ดังภาพที่ 4) และความแม่นยำทางสถิติ



ภาพที่ 4: กระบวนการวิจัยและความคลาดเคลื่อนที่อาจเกิดขึ้นได้ในการวิจัยทางด้านสรีรวิทยาของการออกกำลังกาย

1) การกำหนดขอบเขตการวิจัย "ปริมาณ/จำนวน" (Delimitations)

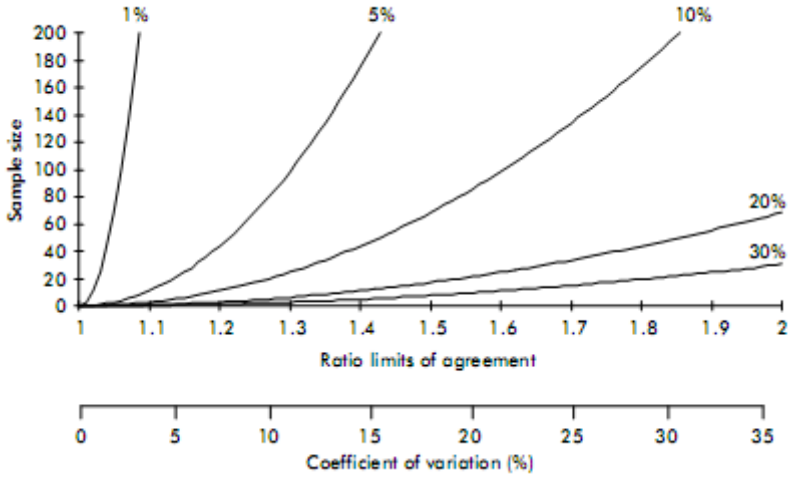
ในกระบวนการวิจัย การเก็บข้อมูลจากกระบวนการวัดควรมีผู้เข้าร่วมทดสอบอย่างน้อย 40 คน หากมีผู้เข้าร่วมน้อยกว่า 40 คน จำเป็นต้องมีการตรวจสอบอย่างละเอียดมากขึ้นของระดับความเชื่อมั่นทางสถิติของค่าความคลาดเคลื่อน เนื่องจากความคลาดเคลื่อนที่ประมาณการจากการคำนวณซึ่งมาจากกลุ่มตัวอย่างที่มีขนาดเล็กนั้นอาจไม่ชัดเจน (Atkinson, 2003)

ในงานวิจัยทางสรีรวิทยาของการออกกำลังกาย การพิจาณาความสอดคล้องกันในลักษณะหรือวิธีการของการวัด เป็นสิ่งจำเป็นที่ต้องให้ความสำคัญ เช่นเดียวกับกระบวนการวัดกับรูปแบบการใช้งานตามแผนของเครื่องมือวัด การพิจารณาถึงความสอดคล้องกันในวิธีการ ที่งานวิจัยทางด้านสรีรวิทยาของการออกกำลังกายควรพิจารณานั่นคือ ประชากรที่เหมือนกัน ช่วงเวลาเดียวกัน ระหว่างกระบวนการทำซ้ำ ๆ โปรโตคอลของการออกกำลังกายหรือรูปแบบของการฝึกการออกกำลังกาย ที่ใกล้เคียงกัน รวมทั้งสภาวะในช่วงของการพักระหว่างการฝึกหรือการออกกำลังกาย ที่สามารถเปรียบเทียบกันได้ในระหว่างกระบวนการทำซ้ำ

การเลือกค่าระดับของความคลาดเคลื่อนที่ถือว่ายอมรับได้ก่อนที่จะศึกษาการเปรียบเทียบระหว่างวิธีการหรือการทำซ้ำ การแบ่งช่วงระดับของความคลาดเคลื่อนนี้ส่วนใหญ่ขึ้นอยู่กับความต้องการของผู้วิจัยในการใช้เครื่องมือวัดผล เพื่อที่สามารถบรรลุวัตถุประสงค์ในการวิจัย เช่น กลุ่มตัวอย่างของผู้มีส่วนร่วมในกระบวนการวัดหรือการทดสอบ หรือการวัดบุคคล เช่น เพื่อวัตถุประสงค์ในการคัดกรองสุขภาพหรือเพื่อสนับสนุนงานด้านวิทยาศาสตร์การกีฬา Atkinson and Nevill (1998) เรียกว่า 'การวิเคราะห์เป้าหมาย' เพื่อบรรลุวัตถุประสงค์ในการวิจัย เป้าหมายของการวิเคราะห์ที่เหมาะสมสำหรับความคลาดเคลื่อนที่เกิดจากการวัด คือ การตั้งค่าที่ดีที่สุดผ่านการคำนวณค่ากำลังทางสถิติ สามารถกำหนดจำนวนหรือขนาดของความคลาดเคลื่อนในการทำซ้ำได้ สามารถอธิบายได้โดยใช้ค่าส่วนเบี่ยงเบนมาตรฐาน ตัวอย่างเช่น บนพื้นฐานของค่ากำลังทางสถิติที่ยอมรับได้ในการตรวจสอบความแตกต่างระหว่างกลุ่มหรือการกำหนดขนาดกลุ่มตัวอย่างที่เหมาะสม (Atkinson and Nevill, 2001) ถ้าจำนวนหรือขนาดของกลุ่มตัวอย่างค่อนข้างใหญ่ ซึ่งอาจเป็นไปได้สำหรับการวิจัยในอนาคตที่สามารถลดผลกระทบต่อการใช้เครื่องมือวัด หรือจากจำนวนหรือขนาดที่กำหนดของความคลาดเคลื่อนในการวัด

สำหรับการใช้เครื่องมือวัดเฉพาะรายบุคคล อาจจำกัดระดับความคลาดเคลื่อนที่ยอมรับได้บนพื้นฐานของความแตกต่างของเฉพาะบุคคลต่อ "สถานการณ์ที่เลวร้ายที่สุด" ซึ่งอาจจะเป็นไปได้ การจำกัด (Bland and Altman, 1999, 2003) เกี่ยวข้องกับการประยุกต์ใช้ค่าทางสถิติของช่วงระดับการยอมรับได้ในความคลาดเคลื่อนตามมาตรฐานที่เกิดจากการวัด (Harvill, 1991) ตัวอย่างเช่น ความแตกต่างที่มีขนาดใหญ่เท่ากับ 5 จังหวะในนาทีที่ 1 ระหว่างการวัดซ้ำสองครั้งของอัตราการเต้นของหัวใจในระหว่างการออกกำลังกาย อาจจะยังคงมีความแตกต่างกันเล็กน้อยที่เกิดขึ้นกับการกำหนดอัตราการเต้นหัวใจให้กับแต่ละบุคคล

การประยุกต์ใช้หลักการง่ายๆ เช่น การใช้ช่วงระดับการยอมรับของความสอดคล้องที่เหมาะสมระหว่างวิธีการวัดหรือการทดสอบ บนพื้นฐานของค่าสัมประสิทธิ์สหสัมพันธ์ที่สูงกว่า 0.9 หรือค่าสัมประสิทธิ์ความแปรปรวน (CV) ที่มีค่าต่ำกว่า 10% เนื่องจากไม่มีความสัมพันธ์ที่เกิดขึ้นระหว่างความคลาดเคลื่อนกับการใช้งานจริงของเครื่องมือวัด (Atkinson, 2003) อย่างไรก็ตาม มีความเห็นบางส่วนของนักวิชาการ ที่กล่าวไว้ว่ากำลังทางสถิติของความคลาดเคลื่อนที่ถูกอ้างถึงจากหลายงานวิจัย ซึ่งได้จากกำลังทางสถิติจากกระบวนการวัดที่หลากหลาย อาจเป็นประโยชน์สำหรับการสร้าง “ระดับความคลาดเคลื่อนที่ยอมรับได้” ที่สามารถนำไปประยุกต์ใช้ได้งานวิจัยทั่วไป ในการเลือกใช้ช่วงระดับความคลาดเคลื่อนที่ยอมรับได้ (Hopkins, 2000)



ภาพที่ 5: รูปโนโมแกรม (nomogram) เพื่อประเมินผลกระทบของระดับความคลาดเคลื่อนในการวัดซ้ำ
ที่มา: Batterham, AM and Aktinson, G. (2005).

จากภาพที่ 5 รูปโนโมแกรม (nomogram) จะถูกนำมาใช้เพื่อประเมินผลกระทบของระดับความคลาดเคลื่อนในการวัดซ้ำว่า "เป้าหมายของการวิเคราะห์" สามารถทำได้หรือไม่ ในการวิจัยด้านสรีรวิทยาของการออกกำลังกาย ค่ากำลังทางสถิติคือที่ระดับ 90% เส้นที่แตกต่างกันแสดงถึงการเปลี่ยนแปลงที่ในค่าระดับ 1%, 5%, 10%, 20% และ 30% เนื่องจากเชิงสมมุติฐานมีการแทรกแซง กำลังทางสถิติของระดับความคลาดเคลื่อนในการวัดที่สามารถนำไปใช้ได้คือ ระดับความสอดคล้อง (LOA) และ ค่าสัมประสิทธิ์ความแปรปรวน (CV) ตัวอย่างเช่นเครื่องมือวัดทางสรีรวิทยาที่มีค่า CV ของการวัดซ้ำได้มีค่าเท่ากับ 5% โดยค่าระดับความคลาดเคลื่อนที่ยอมรับได้ของการตรวจสอบการเปลี่ยนแปลงการทดสอบก่อนและหลังได้เท่ากับ 5% (ใช้ Paired t-test สำหรับการวิเคราะห์ข้อมูล) และมีขนาดตัวอย่างที่เป็นไปได้ (~ 20 คน)

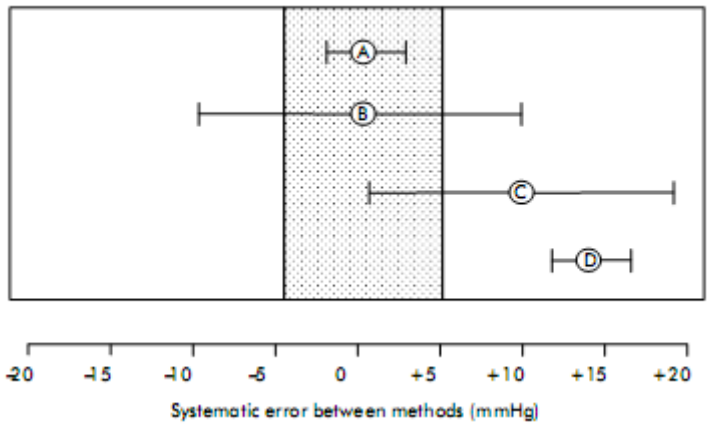
2) การตรวจสอบความคลาดเคลื่อนเชิงระบบ (Systematic Error Examination)

ความคลาดเคลื่อนเชิงระบบเป็นความคลาดเคลื่อนที่เกิดจากความลำเอียง ซึ่งมีโอกาสเกิดขึ้นได้ในทุกขั้นตอนของการวิจัย และมีโอกาสเกิดขึ้นได้เท่ากันทุกกลุ่มประชากร ความคลาดเคลื่อนเชิงระบบอาจเกิดขึ้นจากตัวผู้วิจัยเองโดยความตั้งใจหรือไม่ตั้งใจก็ตาม แต่ความลำเอียงนั้นได้ส่งผลต่อข้อสรุปของการวิจัยยิ่งกว่านั้นความคลาดเคลื่อนเชิงระบบ อาจเป็นผลมาจากข้อบกพร่องของเครื่องมือวัด หรือเกิดจากสภาพแวดล้อมที่ทำการวัดหรือการทดสอบ ความคลาดเคลื่อนในเครื่องมือวัด เป็นผลมาจากความคลาดเคลื่อนที่อาจมาจากชิ้นส่วนทางกายภาพของเครื่องมือวัด เช่น กลไกที่เสื่อมหรือฝืด ความผิดปกติของบางชิ้นส่วน หรือคุณสมบัติของวัสดุประกอบบางตัวที่เสื่อมสภาพ ดังนั้น ผู้วิจัยหรือผู้ทำการวัดจำเป็นต้องให้ความสำคัญกับการเลือกเครื่องมือวัดที่เหมาะสม การปรับแต่งเพื่อเทียบกับค่ามาตรฐาน การวัดซ้ำเพื่อทำการสอบเทียบกับค่า Gold Standard (หรือการตรวจมาตรฐานสูงสุด) หรือ การเพิ่มค่าน้ำหนัก (หรือ Factors) เพื่อชดเชยค่าที่เบี่ยงเบนไป นอกจากนี้ ในส่วนของความคลาดเคลื่อนจากสิ่งแวดล้อม เป็นผลมาจากการเปลี่ยนแปลงหรือความแปรปรวนจากภายนอก

อาทิ การเปลี่ยนแปลงในอุณหภูมิ ความชื้น เป็นต้น ซึ่งผู้วิจัยหรือผู้วัด ควรหาแนวทางหรือวิธีการที่สามารถควบคุมปัจจัยที่มีความไวต่อการวัดทางสรีรวิทยาได้

การเปรียบเทียบความแตกต่างระหว่างวิธีการวัดหรือวิธีการทดสอบ โดยอยู่บนพื้นฐานของระดับที่ยอมรับได้ตามที่กำหนดไว้ล่วงหน้า (ภาพที่ 6 เกี่ยวกับการประยุกต์ใช้ขอบเขตค่าอ้างอิงของระดับความเชื่อมั่นในการตีความของความแตกต่างระหว่างวิธีการวัด)

หากมีความคลาดเคลื่อนเชิงระบบระหว่างกระบวนการทดสอบซ้ำโดยผู้วิจัยใช้วิธีการเดียวกัน เช่น การศึกษาความสามารถในการทำซ้ำ เป็นต้น ผู้วิจัยจำเป็นต้องตั้งข้อสังเกตในกระบวนการการออกแบบของการศึกษาความสามารถในการทำซ้ำ ซึ่งบางทีอาจจะมีผลกระทบจากการตรวจวัดก่อนหน้านี้ที่ใกล้เคียงกับค่าที่ได้จากการวัดที่ได้จริง สถานการณ์ดังกล่าวอาจเกิดขึ้นพร้อมกับการวัดอุณหภูมิภายในห้องตัวอย่าง (Atkinson et al., 2005a) เช่น หากมีการวัดสมรรถนะในรูปแบบการฝึกการออกกำลังกายแล้วพบความแตกต่างเชิงระบบระหว่างการทำซ้ำ การศึกษาความสามารถในการทำซ้ำอาจเกิดขึ้นเนื่องจากผลจากการเรียนรู้ก่อนหน้านี้ ข้อมูลเหล่านี้มีความสำคัญเพื่อเป็นข้อเสนอแนะแก่นักวิจัยในอนาคตได้ ซึ่งอาจต้องอาศัยความคุ้นเคยกับการบันทึกค่าต่าง ๆ ทางสรีรวิทยาก่อนการบันทึกอย่างเป็นทางการด้วย



ภาพที่ 6: เกี่ยวกับการประยุกต์ใช้ขอบเขตค่าอ้างอิงของระดับความเชื่อมั่นในการตีความของความแตกต่างระหว่างวิธีการวัด

ที่มา: Atkinson, G. , Davison, R.C.R. และ Nevill, A.M. (2005).

จากภาพที่ 6 แสดงให้เห็นถึงการประยุกต์ใช้ช่วงระดับของความเชื่อมั่น และ 'ขอบเขตของความสอดคล้อง' (พื้นที่สีเทา) สำหรับความคลาดเคลื่อนระหว่างวิธีการวัดหรือวิธีการทดสอบ การใช้ความสอดคล้องระหว่าง 2 เครื่องวัดความดันโลหิตเป็นตัวอย่าง ถ้าระดับความเชื่อมั่นที่ 95% และตกอยู่ใน "ขอบเขตของความสอดคล้อง" ของ 5 mmHg ที่มีการกำหนดไว้ก่อนหน้านี้ สำหรับค่าความคลาดเคลื่อนระหว่างวิธีการเช่นในกรณีที่เป็นจุด A เราสามารถทำได้ มีเหตุผลพอสมควรว่าความคลาดเคลื่อนที่แท้จริงเชิงระบบระหว่างวิธีการไม่ได้เป็นความสำคัญสำหรับการทดสอบในรูปแบบทางคลินิกหรือทางปฏิบัติ สำหรับจุด D แม้มีระดับความเชื่อมั่นที่ 95% ของ 10.6 mmHg สำหรับความคลาดเคลื่อนเชิงระบบระหว่างวิธีการไม่ได้อยู่ในช่วงขอบเขตของความสอดคล้อง

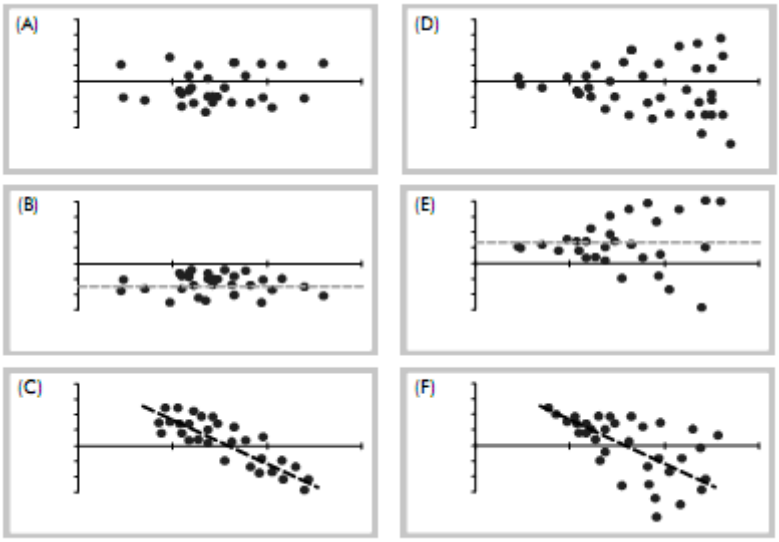
ตามที่กำหนดไว้ ดังนั้นเราจึงควรตรวจสอบให้แน่ใจว่าระดับความคลาดเคลื่อนเชิงระบบจะส่งผลกระทบในทางปฏิบัติ ความกว้างของ CIs สำหรับจุด B และ C อธิบายได้ว่ากลุ่มตัวอย่างหรือประชากรมีความลำเอียง ซึ่งอาจมีความสำคัญในทางปฏิบัติ แต่จะต้องเพิ่มการศึกษาในจำนวน Cases ให้มากขึ้นในกระบวนการทดสอบ หรือการวิจัยให้เหมาะสมกับขนาดที่แท้จริงหรือในกรณีของจุด B อาจมีความคลาดเคลื่อนเชิงระบบในกลุ่มประชากร

การตรวจสอบว่าระดับความคลาดเคลื่อนเชิงระบบมีการเปลี่ยนแปลงไปในช่วงของการวัด ซึ่งสามารถประยุกต์ใช้แผนภาพของ Bland-Altman สำหรับข้อมูลเหล่านี้ได้ (ภาพที่ 6) การปรากฏของ 'สัดส่วน' ของความลำเอียงจะสามารถระบุได้ ถ้าพบจุดบนแผนภาพของ Bland-Altman ที่แสดงถึงแนวโน้มของการลดลงหรือการเพิ่มขึ้นจากช่วงของการวัด ถ้าเกิดลักษณะดังกล่าวนี้ ก็หมายความว่า ความคลาดเคลื่อนเชิงระบบระหว่างวิธีการหรือการทดสอบ มีความแตกต่างกันสำหรับเฉพาะบุคคล ที่จุดสิ้นสุดของช่วงการวัดอยู่ในระดับที่ต่ำและสูง Bland and Altman (1999) และ Atkinson et al. (2005a) มีการถกเถียงถึงประเด็นที่ว่า ความลำเอียงเชิงสัดส่วนนี้สามารถถูกพบและถูกสร้างเป็นแบบจำลองได้อย่างไร

3) การตรวจสอบความคลาดเคลื่อนแบบไม่แน่นอน (Random Error Examination)

ความคลาดเคลื่อนจากความไม่แน่นอน หรือ ความคลาดเคลื่อนแบบสุ่ม อาจกล่าวได้ว่า การที่ไม่สามารถระบุถึงสาเหตุที่แน่นอนได้ เกิดจากการที่ผู้วิจัยดำเนินการวัดในสิ่งเดียวกันในหลาย ๆ ครั้ง แต่ได้ค่าที่แตกต่างกันไป แม้ว่าเครื่องมือวัดถูกปรับแต่งอย่างเหมาะสมแล้วและการใช้งานเครื่องมือวัดหรือวิธีการวัดก็อยู่ภายใต้สภพแวดล้อมที่เหมาะสมต่อการวัดแล้วก็ตาม เป็นความคลาดเคลื่อนที่เกิดขึ้นเองหรือเกิดโดยบังเอิญ ไม่มีทิศทางที่แน่นอน ดังนั้น ผู้วิจัยหรือผู้ดำเนินการวัด ควรเพิ่มจำนวนครั้งของการวัดแล้วหาค่าเฉลี่ยทางสถิติเพื่อที่จะให้ได้ค่าประมาณที่จัดได้ว่าเป็นค่าที่ดีที่สุด หรือผู้วิจัยสามารถกำหนด สูตรในการประมาณค่า แบบแผนการสุ่ม ขนาดของตัวอย่าง

การตรวจสอบระดับความคลาดเคลื่อนแบบไม่แน่นอน ระหว่างวิธีการหรือวิธีการทดสอบที่มีอยู่ (ภาพที่ 7) สถิติที่นำมาใช้เพื่ออธิบายระดับของความคลาดเคลื่อนแบบไม่แน่นอน ได้แก่ Standard Error of Measurement (SEM) (Harvill, 1991) ซึ่งเป็นที่ทราบทั่วกันว่าค่าเบี่ยงเบนมาตรฐาน ค่าสัมประสิทธิ์ความแปรปรวน (CV) ระดับของความสอดคล้อง (LOA) และส่วนเบี่ยงเบนมาตรฐานของความแตกต่าง แต่ละสถิติเหล่านี้จะแตกต่างกันเนื่องจากอยู่บนพื้นฐานปรัชญาที่แตกต่างกัน ตัวอย่างเช่น สถิติ LOA มีรากฐานมาจากงานวิจัยหรือการศึกษาทดสอบและวิเคราะห์ในรูปแบบทางคลินิก Lab และสามารถเห็น ความคลาดเคลื่อนในสถานการณ์ที่เลวร้ายที่สุดที่อาจสังเกตเห็นได้ในตัวบุคคล สถิติ SEM เป็นที่นิยมในหมู่นักจิตวิทยาที่คิดว่าค่าเฉลี่ยของการวัดซ้ำหลายๆครั้งเป็นค่า 'ค่าคะแนนที่แท้จริง' สำหรับแต่ละบุคคล (Harvill, 1991)



ภาพที่ 7: แผนภาพ Bland-Altman ที่แสดงความแตกต่างกันของความสัมพันธ์ระหว่างความคลาดเคลื่อนเชิงระบบและแบบไม่แน่นอนกับขนาดของค่าที่วัดได้

ที่มา: Atkinson, G. , Davison, R.C.R. และ Nevill, A.M. (2005).

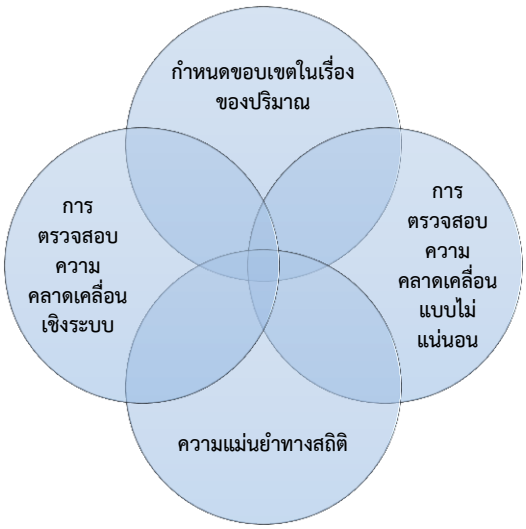
ภาพที่ 7 แสดงตัวอย่างที่แตกต่างกันของความสัมพันธ์ระหว่างความคลาดเคลื่อนเชิงระบบและแบบไม่แน่นอนกับขนาดของค่าที่วัดได้ ดังที่แสดงในแผนภาพ Bland-Altman โดยแผนภาพ (A) ความคลาดเคลื่อนแบบไม่แน่นอนตามสัดส่วน (B) ความคลาดเคลื่อนเชิงระบบซึ่งมีลักษณะเหมือนกัน ความคลาดเคลื่อนแบบไม่แน่นอนยังเหมือนกัน (C) ความคลาดเคลื่อนเชิงระบบซึ่งก็คือ สัดส่วนกับขนาดของค่าที่วัดได้ ชุดความคลาดเคลื่อนแบบไม่แน่นอน (D) ความคลาดเคลื่อนแบบไม่แน่นอนตามสัดส่วนกับความคลาดเคลื่อนเชิงระบบ (E) ความคลาดเคลื่อนที่เป็นรูปแบบเดียวกันกับความคลาดเคลื่อนแบบไม่แน่นอนตามสัดส่วน (F) ความคลาดเคลื่อนเชิงระบบและความคลาดเคลื่อนแบบไม่แน่นอนตามสัดส่วน

การหาความสัมพันธ์ระหว่างความคลาดเคลื่อนแบบไม่แน่นอนและขนาดของค่าที่วัดได้ (Atkinson and Nevill, 1998; Bland and Altman, 1999) หากความคลาดเคลื่อนแบบไม่แน่นอนเพิ่มขึ้นตามขนาดของค่าที่วัดได้แล้ว ควรใช้กำลังทางสถิติอัตราส่วนเพื่ออธิบายความคลาดเคลื่อนในการวัด เช่น ค่าสัมประสิทธิ์ความแปรปรวน (CV) (Nevill and Atkinson, 1998; Bland and Altman, 1999) หากความคลาดเคลื่อนเป็น 'homoscedastic' กล่าวคือมีความสม่ำเสมอในช่วงการวัด ความคลาดเคลื่อนในการวัดสามารถอธิบายได้ในหน่วยการวัดเฉพาะโดยการคำนวณค่า LOA หรือ ค่า SEM ความสัมพันธ์ที่ซับซ้อนระหว่างขนาดของความคลาดเคลื่อนและค่าที่วัดได้ สามารถวิเคราะห์ได้โดยใช้วิธีการที่ไม่ใช่พาราเมตริก อ้างอิงตาม Bland and Altman (1999) หรือโดยการคำนวณค่ากำลังทางสถิติของความคลาดเคลื่อนในการวัดสำหรับตัวอย่างย่อยที่แยกจากกันภายในกลุ่มประชากร (Lord, 1984)

ในการศึกษาความสามารถในการทำซ้ำ ซึ่งเกี่ยวข้องกับการทดสอบซ้ำหลายครั้ง ควรตรวจสอบว่ามีการเปลี่ยนแปลงในความคลาดเคลื่อนแบบไม่แน่นอนระหว่างการทดสอบ โดยดำเนินการทดสอบที่แยกต่างหากและทดสอบแบบซ่อมใหม่หรือไม่ นักวิจัยสามารถสำรวจพบว่าความคลาดเคลื่อนแบบไม่แน่นอนลดลงเมื่อมีการทดสอบหรือการวัดเพิ่มเติม (Nevill and Atkinson, 1998) หากเป็นเช่นนั้นและสอดคล้องกับคำแนะนำข้างต้นสำหรับการสำรวจความคลาดเคลื่อนเชิงระบบ ผู้วิจัยควรที่จะถ่ายทอดหรือให้ข้อเสนอแนะถึงผลที่ได้จากการเรียนรู้นี้ ถึงในกรณีความคลาดเคลื่อนแบบไม่แน่นอนเพื่อให้ผู้ใช้ในอนาคตทราบว่าขนาดของการวัดซ้ำควรมีมากขนาดไหน ซึ่งจะจำเป็นสำหรับการลดความแปรปรวนของความคลาดเคลื่อนทั้งหมดสำหรับการวิจัยหรือการศึกษาได้ในอนาคต

4) ความแม่นยำทางสถิติ (Statistical Precision)

การหาค่าระดับความเชื่อมั่น 95% (CI) สำหรับค่าเฉลี่ยของความคลาดเคลื่อนระหว่างวิธีการวัดหรือวิธีการทดสอบ (Jones et al., 1996) และเมื่อเปรียบเทียบค่า CI นี้กับ 'ขอบเขตความสอดคล้อง' สำหรับสองวิธีของการวัด (รูปที่ 3) ค่าช่วงระดับความเชื่อมั่น CI นี้ไม่เหมือนกับค่ากำลังทางสถิติ LOA การตรวจสอบขอบเขตด้านล่างและด้านบนของ CI นี้ ไม่ควรเปลี่ยนข้อสรุปที่ได้รับเกี่ยวกับการยอมรับความคลาดเคลื่อนของระบบระหว่างวิธีการหรือการทดสอบ (รูปที่ 3) ตัวอย่างเช่น อาจสังเกตความแตกต่างเฉลี่ย 10 mmHg ระหว่างอุปกรณ์วัดความดันโลหิต ที่ระดับ 95% CI อาจเป็น 4 ถึง 24 mmHg ซึ่งหมายความว่า ความแตกต่างเฉลี่ยของกลุ่มประชากรระหว่างวิธีการอาจจะมากถึง 4 mmHg ในทิศทางเดียวหรือมากถึง 24 mmHg ในอีกทางหนึ่งเฉพาะช่วงระดับค่า CI ที่แคบขึ้น โดยส่วนใหญ่จะมีขนาดตัวอย่างมากขึ้น จะช่วยให้สามารถสรุปเกี่ยวกับความคลาดเคลื่อนของระบบได้มากขึ้น Atkinson and Nevill (2001) และ Atkinson et al. (2005a) กล่าวถึงการใช้ระดับ CI และข้อจำกัดเพิ่มเติม การคำนวณค่าระดับความเชื่อมั่นสำหรับค่าสถิติของความคลาดเคลื่อนแบบไม่แน่นอน ดังที่กล่าวมาข้างต้น การตรวจสอบข้อเท็จจริงของระดับ CI ไม่ควรเปลี่ยนแปลงการตัดสินใจที่ได้ทำขึ้นเกี่ยวกับการยอมรับความคลาดเคลื่อนแบบไม่แน่นอน ตัวอย่างเช่นวิธีการวัดทางสรีรวิทยาที่มีค่า CV ในการทำซ้ำได้ถึง 30% และค่า CI ที่เกี่ยวข้องที่ระดับค่า 25-35% แสดงถึงระดับความสามารถของการทำซ้ำที่ไม่ดีแม้จะมีการพิจารณาขอบเขตระดับล่างของค่า CI ดังนี้ Bland and Altman (1999) ได้ให้รายละเอียดเกี่ยวกับขอบเขตของความสอดคล้อง Hopkins (2000) แสดงวิธีการคำนวณขอบเขตของระดับความเชื่อมั่นสำหรับ CV และ Morrow and Jackson (1993) ได้ให้รายละเอียดสำหรับความสัมพันธ์แบบ Intraclass โดยสถิติที่ใช้วัดการสอดคล้องและเป็นที่นิยมใช้ คือ สัมประสิทธิ์สหสัมพันธ์ภายในชั้น (Intraclass Correlation Coefficient : ICC) การวัดความสอดคล้อง (Agreement Measurement) ถูกใช้อย่างแพร่หลายเพื่อพิจารณาถึงสิ่งที่วัดสิ่งเดียวกันโดยผู้วัดหลายคนหรือเครื่องมือวัดหลายชนิดให้ผลการวัดได้สอดคล้องกัน ในปัจจุบันนี้ สัมประสิทธิ์สหสัมพันธ์ภายในชั้นถูกนำมาประยุกต์ใช้กันมากขึ้นในวงการเพื่อประเมินผลความน่าเชื่อถือระหว่างผู้สังเกต (Inter-rater Reliability) ความน่าเชื่อถือภายในผู้ประเมิน (Intra-Rater Reliability) และการทดสอบและทดสอบซ้ำ (Test-Retest Reliability) ดังนั้นการเลือกใช้แบบจำลองเพื่อใช้ในการคำนวณค่าสัมประสิทธิ์สหสัมพันธ์ภายในจำเป็นต้องได้รับการพิจารณา และควรต้องเลือกแบบจำลองสำหรับการวิเคราะห์ให้เหมาะสมเพื่อความถูกต้องของการวิเคราะห์



ภาพที่ 8: แนวทางในการตรวจสอบในการประเมินเครื่องมือวัดสำหรับการวิจัยทางสรีรวิทยาของการออกกำลังกาย

บทสรุป

จากบทความนี้ได้อธิบายถึงองค์ประกอบของรายการตรวจสอบที่มีความสำคัญสำหรับนักสรีรวิทยา (Physiologists) ด้านการออกกำลังกาย ที่มีความสนใจในการตรวจสอบประสิทธิภาพการทำงานของกระบวนการวัดผลหรือเครื่องมือวัดที่มีลักษณะเฉพาะ ปัญหาที่สำคัญที่สุดที่ต้องพิจารณาโดยทั่วไป ได้แก่ การใช้เครื่องมือวัด การศึกษาระดับของความคลาดเคลื่อนเชิงระบบและความคลาดเคลื่อนแบบไม่แน่นอนระหว่างวิธีการวัดหรือกระบวนการทดสอบซ้ำ และค่าระดับความแม่นยำทางสถิติที่เหมาะสมสำหรับการคำนวณค่าประมาณการของค่าความคลาดเคลื่อนเหล่านี้ ปัญหาทั้งหมดเหล่านี้ไม่สามารถดำเนินการได้เพียงแค่ประยุกต์ใช้ในหลักการคำนวณกำลังทางสถิติเพียงค่าเดียวได้ ดังนั้นนักสรีรวิทยาการออกกำลังกายควรตระหนักถึงค่ากำลังทางสถิติต่างๆ ซึ่งใช้ในการวัดระดับความสอดคล้องและระดับของความคลาดเคลื่อนโดยเฉพาะอย่างยิ่งในแง่มุมของระดับความคลาดเคลื่อนที่สามารถส่งผลกระทบต่อความถูกต้องของข้อสรุปจากการศึกษาหรือการวิจัยได้ ประเด็นหนึ่งที่น่าสังเกตคือ มีหลายองค์กรวิชาชีพ ที่มีแนวทางที่สามารถใช้อ้างอิงเป็นหลักฐานเชิงประจักษ์ที่กล่าวถึงเกณฑ์ของค่าระดับความสอดคล้องที่ยอมรับได้ในวิธีการวัด และในวิธีการวัดที่เกี่ยวข้องกับค่าเฉลี่ยและค่าส่วนเบี่ยงเบนมาตรฐานของความแตกต่าง (O'Brien, 1998) แต่น่าเสียดายที่ค่ามาตรฐานที่ได้รับการยอมรับเหล่านี้ ค่อนข้างหายากที่จะถูกนำมาใช้อ้างอิงในงานวิจัยหรือการทดสอบทางสรีรวิทยาการออกกำลังกาย

References

Atkinson, G. (2003). What is this thing called measurement error? In T. Reilly, M.Marfell-Jones (eds), Kinanthropometry VIII: Proceedings of the 8th International Conference of the International Society for the Advancement of Kinanthropometry (ISAK) pp. 3–14. London: Taylor and Francis.

Atkinson, G. and Nevill, A.M. (1998). Statistical methods for assessing measurement error (reliability) in variable relevant to sports medicine. Sport Medicine, 26, 217-238.

Atkinson, G. and Nevill, A.M. (2001). Selected issues in the design and analysis of sport performance research. Journal of Sports Sciences, 19: 811–827.

Atkinson, G., Davison, R.C.R. and Nevill, A.M. (2005a). Performance characteristics of gas analysis systems: what we know and what we need to know. International Journal of Sports Medicine, 26 (Suppl. 1): S2–S10.

Bland, J.M. and Altman, D.G. (1999). Measuring agreement in method comparison studies. Statistical Methods in Medical Research, 8: 135–160.

Bland, J.M. and Altman, D.G. (2003). Applying the right statistics: analyses of measurement studies. Ultrasound and Obstetrics in Gynecology, 22: 85–93.

Bird, S. and Davidson, R. (1997). Physiological Testing Guildelines, 3rd edn. Leeds: The British Association of Sport and Exercise Sciences.

Harvill, L.M. (1991). An NCME instructional module on standard error of measurement. Educational Measurement: Issues and Practice, 10: 33–41.

Hopkins, W. (2000). Measures of reliability in sports medicine and science. Sports Medicine, 30: 1–15.

Jones, B. et al. (1996). Trials to assess equivalence: the importance of rigorous methods. British Medical Journal, 313: 36–39.

Lord, F.M. (1984). Standard errors of measurement at different ability levels. Journal of Educational Measurement, 21: 239–243.

Morrow, J.R. and Jackson, A.W. (1993). How ‘significant’ is your reliability? Research Quarterly for Exercise and Sport, 64: 352–355.

Nevill, A.M. and Atkinson, G. (1998). Assessing measurement agreement (repeatability) between 3 or more trials. Journal of Sports Sciences, 16: 29.

O’Brien, E. (1998). Automated blood pressure measurement: state of the market in 1998 and the need for an international validation protocol for blood pressure measuring devices. Blood Pressure Monitoring, 3: 205–211.