SPARTA SCIENCE

Balance Assessment Overview

By Greg Olsen, PhD and Jessica Hinman, PhD

BACKGROUND

Balance is a central and complex aspect of human movement - it requires input from multiple sensory systems, musculoskeletal responses and coordination by the brain and spinal cord.

People desire “good balance” for obvious reasons including reduced risk of falling and improved athletic and day-to-day physical performance. Balance capabilities are not static nor inherent. They can be improved through exercise and can be adversely impacted by a variety of different underlying physical ailments, including:

- Musculoskeletal injuries or deficits
- Brain injuries & neurological disorders
- Sensory disorders

The testing of balance ability, then, is an important evaluation tool that can guide health improvement and medical interventions when necessary. Clinical testing by health professionals has been used in various forms for more than a century. Common practice is dominated by what are typically referred to as ‘**functional tests of balance**’. Examples, listed in Appendix A, include the Romberg Test, the BESS Test, and the TUG (Timed Up & Go) Test.

Functional tests are useful - particularly for identifying severe balance abnormalities. They have significant drawbacks, however, that include requirements for skilled clinicians, limited sensitivity and specificity of measurement, and test duration. These drawbacks limit the ability to differentiate among different balance issues or to capture small differences in balance ability [1].

More recently, assessment approaches based on the capture of high-bandwidth time series data have come to the forefront. These methods, originally developed in the research community, are entering mainstream use, fueled by the rapidly expanding availability of computing technology and data.

In these time series based approaches, the goal is to capture the complexity of a person’s response to some form of balance challenge through sensor data. Data is collected continuously over a prescribed period of time - typically thousands of individual measurements per second or more from multiple sensors. Example sensing approaches include:

- Video capture systems
- Body-attached inertial sensors
- In-shoe pressure sensitive inserts
- Force plates

Differentiating movement features or patterns, then, are extracted from the collected kinetic or kinematic data which can be used directly or in computational models to characterize a person’s balance.

This type of balance assessment offers many advantages over traditional functional tests, the most important being the ability to capture fine-grained objective data and to scale testing to population scale. The following sections of this paper explore more details of balance assessment based on time series data analysis.

CONSIDERATIONS IN TIME SERIES BASED ASSESSMENTS

While examining different potential approaches to assessment, there are several relevant considerations, including:

- **Feasibility:** How easy is it to assess people? What is required in terms of time, supervision, clinical expertise or apparatus? How challenging are the instructions the subject must follow?
- **Scale:** Can an assessment be consistently executed at sufficient population scale to reap the benefits of “big data” analysis?
- **Reliability:** How well does a set of measurements differentiate people from each other?
- **Usefulness:** Are the measurements useful as predictors of health conditions of interest? Do they work for many different health conditions? Do the measurements lead to more effective treatments?

Passive, wearable-based approaches (e.g. phone in pocket, wristband or ring-based sensors worn over extended periods of time) are appealing in their ability to do data collection in the background. While they can provide some insights into balance, these passive approaches provide an insufficient foundation for scalable balance assessment. Limitations include:

- Required provisioning of wearable devices
- Available scope of captured sensor data
- Comparability across device types, manufacturers, environmental conditions and placement on individuals

Balance assessment quality benefits from an ***active*** test, where the subject follows prescribed movements in a set period of time. The movements can be simple, but they in some way create a repeatable balance response that produces reliable metrics with sufficient granularity that they can be usable in prediction models. Of the various approaches to capturing time series data in an active test, force plate-based approaches currently offer the best path to scalable and information rich balance assessment.

Balance assessments on a force plate can be executed in one or two minutes, with minimal or no supervision on a simple apparatus while still producing information-rich signal data. They can be used very broadly across the population. The remainder of this document, then, explores details of time series-based balance assessment using a force plate. Appendix B provides some discussion of alternative sensor approaches.

SPARTA SCIENCE

FORCE-PLATE BASED BALANCE TESTING

In a balance assessment, force plates measure ground reaction forces of an upright person attempting to balance [2]. The figures below show the basic physics.

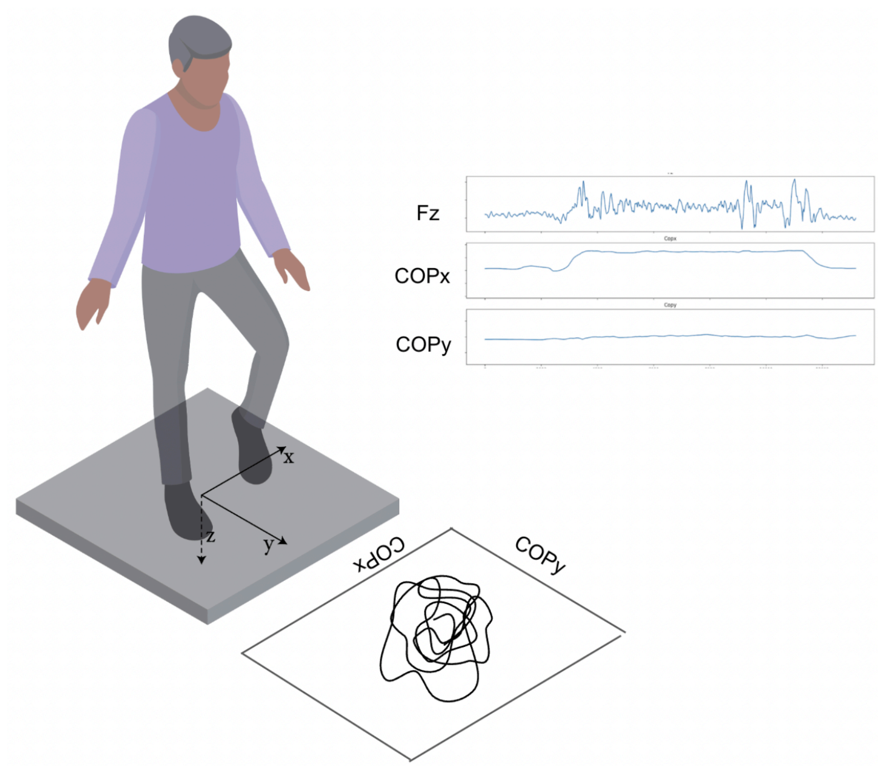

Figure 1

The force plate resolves the sensor data into ***time series*** representations of resultant vertical force (Fz) and center of pressure measures in the plane of the plate (COPx and COPy) representing medial-lateral and anterior-posterior directions respectively.

For assessment, a person executes a protocol (a guided sequence of timed movements) and the time series data is captured. This protocol can consist of sequences composed with basic elements (with variable timing) such as the following:

- Stand still on both feet shoulder width apart
- Raise/lower one foot

Protocols can also specify variations such as:

- Eyes closed vs. open
- Footwear vs. barefoot

Test protocols can be adapted to the abilities of the population being assessed and to the goals of the assessment. In general, however, minimizing variations in protocol or aspects of protocol that reduce participant compliance is important to achieving the maximum scale of comparable test results. Therefore, the following minimal set is preferred:

- Alternating single leg raises with shoes on
- Two legs with shoes on (for individuals incapable of balancing on a single leg)

After a test protocol is executed, analysis starts with transforming the millions of individual Fz, COPx, & COPy measurements into a set of distinguishing features and using those features to characterize the balance ability of the subject.

EXTRACTING BIOMARKERS FROM THE DATA

Here we use the terms biomarker and feature almost interchangeably to describe metrics extracted from the time series data for purposes of analysis. Feature is the term used generically in machine learning contexts and biomarker connotes that the metric reflects a person's health state in some manner.

FEATURES FROM TIME SERIES DATA

Figure 2 below shows the traces for a section of a sample balance test where the user executed a protocol where they started on two feet, raised their left foot and then lowered that foot to restore a two-foot balance position.

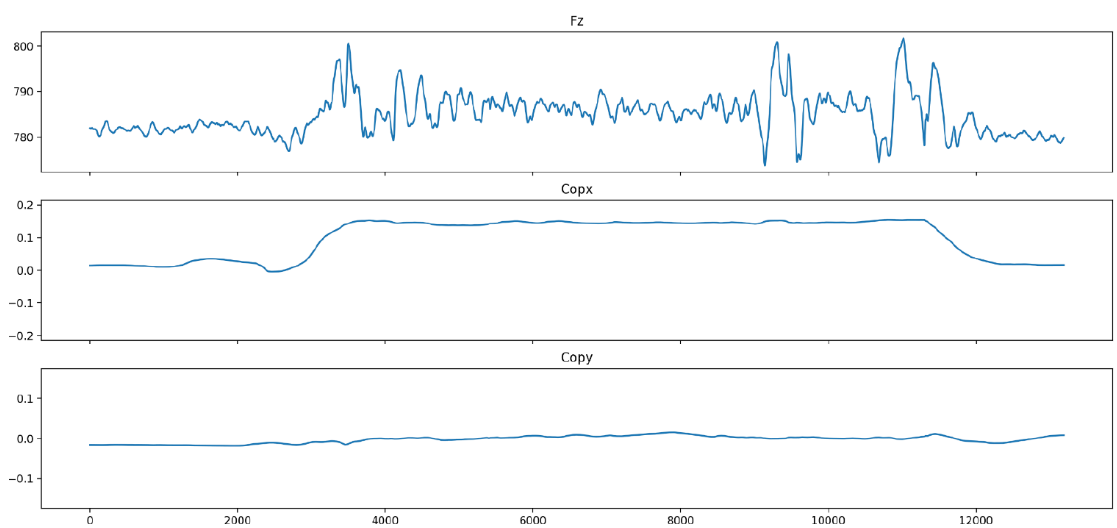

Figure 2

Many choices could be made as to how to extract features from the three signals. For example, we could choose to first isolate different comparable sections of the test like the on-two-feet periods, the on-one-foot period and the transitions (2 foot to 1 foot transition and the 1 foot to 2 foot). For these three signals, there is a substantial amount of information available for evaluation of the person's balance response.

Once segments in the time series data are identified, there are a variety of methods of extracting biomarkers from the signal data in those segments, including:

- Biomechanics/physics-based calculations
- Signal analysis oriented calculations
- Unsupervised machine learning techniques
- Supervised machine learning techniques

BIOMECHANICS/PHYSICS-BASED CALCULATIONS

The most common approach to biomarker extraction from time series movement data has been to calculate quantities that are based on the physics of the movement being tested (e.g. averages or peak values of displacements, velocities, accelerations, forces). In the case of balance testing examples include:

- Average velocity measures for the center of pressure (or sway velocity)
  - Resultant magnitude
  - Anterior-posterior
  - Medial-lateral
- Displacement measures
- Transition time measures

SIGNAL ANALYSIS ORIENTED CALCULATIONS

Additional biomarkers can be extracted by applying signal processing techniques to the data that look for things like information content, spectral properties, or temporal correlations. Examples of quantities that could be used for biomarkers include:

- Entropy measures (e.g. multi-scale, multivariable sample entropy)
- Spectral energy measures (energy in different bands of the frequency spectrum)
- Phase relationship measures (e.g. vector coding, continuous relative phase)

UNSUPERVISED MACHINE LEARNING

A variety of unsupervised machine learning techniques can also be used to generate high quality biomarkers. Example approaches include:

- Direct extraction of biomarkers from patterns present in the raw signal data using autoencoders and temporal convolutional neural networks.
- Dimension reduction and optimization techniques applied to sets of features produced through other means (e.g. physics and signal processing based calculations) to produce derived biomarkers optimized for reliability and independence.
- Clustering techniques to identify significant categorical groups based on sets of features produced through other means.

SUPERVISED MACHINE LEARNING

Where labeled data (i.e. ground truth information for conditions or performance capabilities) is available, supervised machine learning techniques can be used to produce biomarkers as the output of trained prediction models. Potential sources of label data include:

- Recorded injury or fall events
- Other test or survey results (e.g. risk or condition assessment surveys, functional tests, performance tests)
- Insurance claims data
- Occupational availability & productivity data

Example types of biomarkers that can be produced from trained models of this type include:
- Injury risk scores
- Fall risk scores
- Athletic performance scores
- Predicted values for other tests

SYMMETRY BIOMARKERS

In many balance testing protocols, separate left and right side sections are assessed - enabling additional evaluations of symmetry for extracted features.

TRIALS & SCANS

In order to improve the reliability of measurements, most test protocols involve some repetition and error checking of individual ***trials*** in order to complete an individual ***scan*** or assessment. Repetition is used in two ways:

- Automated exclusion of 'invalid' trials: 'Bad' tests can occur due to non-compliance with a specified test protocol or other issues such as inadvertent contact with the plate. Test software can detect many of these events and can direct a repeat of the test.
- Averaging of metrics across successful trials: Incorporating a single repetition the movement and averaging extracted metrics can significantly improve measurement reliability.

USING BIOMARKERS TO ASSESS BALANCE CAPABILITIES

Clearly, a large quantity of information can be extracted from a simple force plate-based balance assessment through many distinct biomarkers. Patterns in this data have been shown to reveal insights into many health concerns related to balance, including:

- Falls
- MSK injuries
- Sensory issues such as vestibular disorders
- Traumatic brain injury
- Chronic neurological disorders such as Parkinson's Disease

Many individual biomarkers have been the subject of clinical research studies, e.g.:

- Lower sway velocities have been associated with better health. Studies have shown higher sway velocities to be correlated with increased fall risk and post-concussion symptoms [3].
- Higher entropy values have been associated with better health. Measurement of the 'complexity' of dynamic physiological signals is an increasingly important assessment tool [4-6]. Higher complexity implies a more well adapted balance control capability. The intuition is that higher complexity of response implies better ability to make fine adjustments at different timescales in order to maintain stability in the face of unpredictable destabilizing perturbations [7].
- Vector coding is frequently used to represent gait data. Vectors are created between data points on relative motion, or angle-angle, plots; these vectors can then be used to assess and analyze coordination [8].
- Continuous Relative Phase is an alternative measure of coordination between two joints or segments that is often employed in gait analysis. In addition to assessments of coordination while walking or running, it has been used to differentiate between healthy individuals and those with movement or neurological disorders [9,10].

Beyond research contexts, however, clinical use of balance testing-based biomarkers have been limited - in the scale of testing, the number of biomarkers, and the diagnostic uses. Where balance testing has been deployed, a single biomarker (most often sway velocity) or a small set of biomarkers are used. The use of an expanded set of biomarkers and large scale testing and data collection enable enormous improvements in balance diagnostic capabilities.

Balance is a multifaceted response that relies on neurological, sensory, and musculoskeletal interactions, and it cannot be easily characterized through a small number of biomarkers. To achieve a better understanding we need:

- Sufficient feature granularity to expose important differences among individuals and to provide a path for explaining the sources of those differences.
- Sufficient cohort data volume to provide a strong basis for comparison.
- Application of machine learning techniques to pragmatically map patterns in balance response data to practical metrics, e.g.:
  - Fall risk scores
  - Injury risk scores
  - Treatment progress measures

SUMMARY

Balance assessment is a rapidly developing practice with implications for a wide variety of fields, including athletic training, rehabilitation, healthy aging, and clinical diagnostics and treatment. Recent advancements in measurement technology and analytic techniques have vastly broadened the realm of possibilities for collecting and evaluating balance data. As best practices emerge, there is increasing potential for individualized balance assessment leading to tailored guidance for performance improvement, injury prevention, and targeted therapy and rehabilitation.

APPENDIX A: FUNCTIONAL TESTS OF BALANCE

A variety of ‘functional tests’ have been used by health professionals to assess an individual’s postural control. These tests generally specify a basic balance related movement and use a simple measurement (such as a time or a distance) or a subjective score assessed by a clinician. Examples include:

- Romberg Test: evaluates the impact of loss of vision (via closed eyes) on balance in a standing position. In this commonly-used test, a patient stands with their feet close together, arms near their sides, and eyes open while the administrator of the test monitors and notes the degree and position of any swaying. The patient then closes their eyes, and the administrator again notes the degree and position of swaying, which is compared with that observed during the eyes-open phase. A patient demonstrating significantly greater sway or imbalance during the eyes-closed phase of the test is considered to have tested positive for Romberg’s sign [11].
- Functional Reach Test: measures the distance beyond arm length a person can reach while maintaining a fixed standing position [12].
- Berg Balance Scale: a clinician-scored assessment of 14 functional activities, including sitting and standing [13,14].
- Performance-Oriented Mobility Assessment (POMA): a clinician-scored assessment of balance across 14 items and gait across 10 items [15].
- Timed Up and Go Test: measures the duration of a series of sequential functional tasks [16].
- Star Excursion Test: evaluates ability to stand on one leg while reaching in eight directions, separated by 45 degrees each, with the other leg [17,18].
- BESS: The BESS (Balance Error Scoring System): assesses deviations from three different stances on firm and soft surfaces with eyes closed [19,20].
- BESTest & Mini-BESTest: a clinician-scored test to differentiate balance into 6 underlying systems that may constrain balance [21].

APPENDIX B: TIME SERIES BALANCE ASSESSMENT

As mentioned previously, there are three primary approaches to collecting data for balance assessment: video-based motion capture of subject, on-body inertial sensors, and ground reaction force measurement.

VIDEO-BASED MOTION CAPTURE

Video is an increasingly convenient tool for kinematic analysis, as video can potentially be captured by subjects with their own phones and no specific expertise. A typical setup for accurate, consistent motion capture requires a lab setting with expensive equipment, setup, and data processing effort. As camera technology on mobile phones has advanced, more mass market approaches to transforming video into 3D representations of human movement are starting to be developed [22, 23]. For assessing balance, however, video-based approaches have limitations with respect to the precision and scope of what they can measure. Balance response involves many small amplitude and high frequency micro-adjustments that cannot be accurately extracted from video in a feasible manner. Video, however, can be a useful complementary tool to other approaches in its ability to document movements of the subject during testing.

ON-BODY INERTIAL SENSORS

Like video technology, accelerometer tech has become commoditized to the point that every smartphone contains an IMU (inertial measurement unit) that can accurately measure 6 degrees of freedom. However, because balance response involves complex small movements of all parts of the body, accurate assessments would require IMUs placed at many separate locations on the body (typical rigid body models of the human body include 15 distinct segments). The challenges posed in instrumentation, data collection, and very high dimensional data analysis limit the practicality of this approach. Like video, limited use of accelerometers can be complementary to other analyses.

GROUND REACTION FORCES

A third approach involves measuring the force exerted onto the ground by the subject while performing a balance movement. Understanding how a subject adjusts the force they exert on the ground in order to maintain balance provides the basis of analysis. The primary method of data collection is the use of a force plate. A force plate is basically a high-tech scale that can dynamically measure vertical force and how that force moves on the plane of the plate surface at high resolution.

Pressure sensing shoe inserts can also be used to characterize force between the feet and ground. Inserts can provide some unique insights (e.g. pressure distribution across the sole of the foot), but present challenges in terms of reliable metric comparisons across different subjects as well as testing pragmatic difficulties at scale (e.g. foot geometry variation, alignment of insert within the shoe).

REFERENCES

1. Mancini M, Horak FB. The relevance of clinical balance assessment tools to differentiate balance deficits. Eur J Phys Rehabil Med. 2010 Jun;46(2):239–48.
2. Chen,B.;Liu,P.;Xiao,F.; Liu, Z.; Wang, Y. Review of the Upright Balance Assessment Based on the Force Plate. Int. J. Environ. Res. Public Health 2021, 18, 2696. https:// doi.org/ 10.3390/ijerph18052696
3. Zhou, J., Habtemariam, D., Iloputaife, I., Lipsitz, L.A., Manor, B., 2017. The Complexity of Standing Postural Sway Associates with Future Falls in Community-Dwelling Older Adults: The MOBILIZE Boston Study. Sci Rep 7, 2924. https://doi.org/10.1038/s41598-017-03422-4
4. Stergiou, N. (2016). Nonlinear Analysis for Human Movement Variability (1st ed.). CRC Press. https://doi.org/10.1201/9781315370651
5. Costa M, Goldberger AL, Peng CK. Multiscale entropy analysis of complex physiologic time series. Phys Rev Lett. 2002 Aug 5;89(6):068102. doi: 10.1103/PhysRevLett.89.068102. Epub 2002 Jul 19. PMID: 12190613.

6. Busa, M.A., van Emmerik, R.E.A., 2016. Multiscale entropy: A tool for understanding the complexity of postural control. J Sport Health Sci 5, 44–51. https://doi.org/10.1016/j.jshs.2016.01.018
7. Manor B, Costa MD, Hu K, et al. Physiological complexity and system adaptability: evidence from postural control dynamics of older adults. J Appl Physiol (1985). 2010;109(6):1786-1791. doi:10.1152/japplphysiol.00390.2010
8. Stock, H., Furlong, L.-A., Wilson, C., Emmerik, R.V., Preatoni, E., 2018. NEW DEVELOPMENTS IN VECTOR CODING METHODS FOR ASSESSING COORDINATION VARIABILITY. ISBS Proceedings Archive 36, 940.
9. Lukšys, D., Jatužis, D., Jonaitis, G., Griškevičius, J., 2021. Application of continuous relative phase analysis for differentiation of gait in neurodegenerative disease. Biomedical Signal Processing and Control 67, 102558. https://doi.org/10.1016/j.bspc.2021.102558
10. Barela, J.A., Whitall, J., Black, P., Clark, J.E., 2000. An examination of constraints affecting the intralimb coordination of hemiparetic gait. Human Movement Science 19, 251–273. https://doi.org/10.1016/S0167-9457(00)00014-2
11. Lanska DJ, Goetz CG. Romberg's sign: development, adoption, and adaptation in the 19th century. Neurology. 2000 Oct 24;55(8):1201–6.
12. Duncan PW, Weiner DK, Chandler J, Studenski S. Functional reach: a new clinical measure of balance. J Gerontol. 1990 Nov;45(6):M192-197.
13. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health. 1992 Aug;83 Suppl 2:S7-11.
14. Berg K, Norman KE. Functional assessment of balance and gait. Clin Geriatr Med. 1996 Nov;12(4):705–23.
15. Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. J Am Geriatr Soc. 1986 Feb;34(2):119–26
16. Mathias S, Nayak US, Isaacs B. Balance in elderly patients: the "get-up and go" test. Arch Phys Med Rehabil. 1986 Jun;67(6):387–9.
17. Gray, GW. Lower Extremity Functional Profile. Wynn Marketing, Inc.; 1995.
18. Gribble PA, Hertel J, Plisky P. Using the Star Excursion Balance Test to Assess Dynamic Postural-Control Deficits and Outcomes in Lower Extremity Injury: A Literature and Systematic Review. J Athl Train. 2012;47(3):339–57.
19. Bell DR, Guskiewicz KM, Clark MA, Padua DA. Systematic Review of the Balance Error Scoring System. Sports Health. 2011 May;3(3):287–95.
20. Guskiewicz KM. Balance assessment in the management of sport-related concussion. Clin Sports Med. 2011 Jan;30(1):89–102, ix.
21. King LA, Priest KC, Salarian A, Pierce D, Horak FB. Comparing the Mini-BESTest with the Berg Balance Scale to Evaluate Balance Disorders in Parkinson's Disease. Parkinsons Dis. 2012;2012:375419. doi: 10.1155/2012/375419. Epub 2011 Oct 24. PMID: 22135761; PMCID: PMC3202113.
22. Uhlrich, S.D., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A.S., Hicks, J.L., Delp, S.L., 2022. OpenCap: 3D human movement dynamics from smartphone videos. https://doi.org/10.1101/2022.07.07.499061
23. Aoyagi, Y., Yamada, S., Ueda, S., Iseki, C., Kondo, T., Mori, K., Kobayashi, Y., Fukami, T., Hoshimaru, M., Ishikawa, M., Ohta, Y., 2022. Development of Smartphone Application for Markerless Three-Dimensional Motion Capture Based on Deep Learning Model. Sensors (Basel) 22, 5282. https://doi.org/10.3390/s22145282