Verisense Validation According to the V3 Validation Framework

Geoffrey Gill[1] and Matthew R Patterson

***Abstract*— This paper summarizes the validation that has been performed for the Verisense Inertial Measurement Unit (IMU) remote monitoring sensor, which utilizes the open source GGIR analysis platform to process raw acceleration data to obtain clinically meaningful activity and sleep metrics. This paper employs the V3 framework proposed by the Digital Medicine Society and others which includes three levels of validation: verification, analytical validation and clinical validation. This paper provides a practical example of how this framework can be employed.**

I. INTRODUCTION

This paper provides a practical example of the use of a three-level validation approach described as the V3 framework [7] using the Verisense platform. The Verisense platform provides remotely captured biometric data from the patients home and community setting. Verisense hardware is based off the Shimmer platform [1], which has been used extensively in clinical research applications for the past decade. These applications include falls risk [2], bradykinesia [3], Parkinsons Disease [4] and anorexia nervosa [5] to name a few. For sleep and activity detection, Verisense calls upon the widely used GGIR package [6].

The V3 medical device validation framework is considered in this paper [7]. The V3 validation framework was designed by industry leaders and published by the Digital Medicine Society to highlight the various levels of validation required for a digital health tool to be used in clinical research. This consists of three levels; verification, analytical validation and clinical validation. Verification consists of validation of the sample-level data from a sensor. Analytical validation consists of assessing the performance of the algorithm to predict behavioral metrics. Finally, clinical validation validates the measure in the stated context of use. Figure 1 shows the summary of the V3 framework.

II. VERIFICATION

The first stage of the V3 validation process evaluates and demonstrates the performance of a sensor technology at the sample-level data it generates, against a pre-specified set of criteria.

[1]Geoffrey Gill is with Shimmer Research, 810 Memorial Dr., Suite 109, Cambridge, MA 02139, USA `ggill@shimmersensing.com`

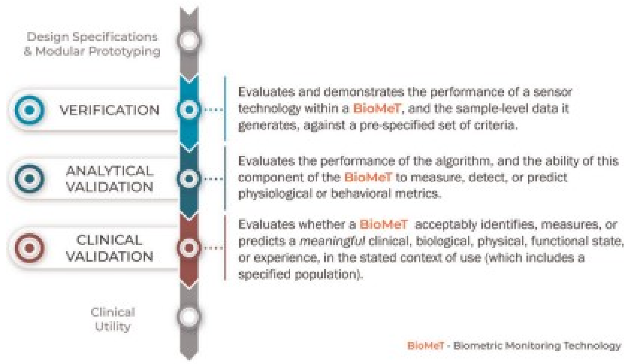

Fig. 1. Summary of the V3 framework. Image used from [7].

A. Epoch-by-Epoch Level Validation

An epoch-by-epoch level validation was carried out on Verisense outputs by researchers at the Letterkenny Institute of Technology [8]. Fifteen adults (11 males, 23.4 ± 3.4 years and 4 females, 29 ± 12.6 years) wore Verisense as well as the reference actigraphy monitor for 48 hours in the free-living protocol. Twelve adults (11 males, 23.4 ± 3.4 years and 1 female, 22 ± 0 years) wore both monitors for the duration of the supervised protocol. Participants performed their regular, at home routine for the free-living protocol. For the supervised protocol, participants performed walking on a treadmill and overground at various speeds as well as ascending and descending stairs. Agreement between the reference actigraph and Verisense was high for both a free living protocol (r = 0.85) as well as a structured protocol (r = 0.78). The conclusion of this study was, ”Verisense, a novel research-grade wearable device, produces activity and sleep parameters that are comparable to a research-grade actigraph”.

This study also demonstrates some of the challenges of performing comparative studies. The reference wearable could not provide raw acceleration data but only activity counts in 15 second intervals. It was therefore necessary to convert the Verisense raw data into a form that would be comparable. The research group chose to use ENMO (Euclidean norm minus one) to correlate with the reference wearable activity count. The correlations were high. Based on these correlations it is clear that ENMO is a good approximation to the activity measure used by the reference device and that Verisense has been independently verified according to the first stage of the V3 validation approach for clinical trials. This is not surprising, both systems use off the shelf accelerometers which have been fully validated and characterized by the manufacturers. If available from the reference device, high correlation with the raw acceleration data would be expected.

B. Accelerometer Auto-Calibration

The Verisense IMU provides raw, tri-axial acceleration data, which is processed via the GGIR software to calculate clinically meaningful sleep and activity metrics. The accelerometer chip on the Verisense IMU provides a linear relationship between the electrical signal and acceleration for the sensor ranges that the chip supports. The noise level in these operating modes range from 1.3-4.5 mg, which are well below a meaningful level for activity and sleep analysis. Acceleration signals are calibrated using the GGIR autocalibratoin method [45]. This works by detecting periods of wear in which the patient is stationary and using the moving average of those periods across each of the accelerometer axes to generate a three-dimensional ellipsoid. Properly calibrated outputs would result in an ellipsoid with a radius of 1 g, so any deviations from this are used to calibrate the accelerometer outputs. This method was shown to provide accurate data in cohorts from the UK (n=921), Kuwait (n=120), Cameroon (n=311) and Brazil (n=200). The authors of this study conclude that, "Results indicate that the autocalibration method works under a wide range of experimental conditions, spanning different geographical latitudes, different seasons affecting temperature variation during the day, different populations affecting movement and activity patterns, different built environments, and different adult age groups" [45].

III. ANALYTICAL VALIDATION

The second validation stage in the V3 process is the analytical validation. The purpose of the analytical evaluation is to evaluate the performance of the algorithm, and the ability of this component of the tool to measure, detect or predict physiological or behavioral metrics. The Verisense system utilizes the GGIR package to process raw accelerometry data into clinically useful activity and sleep metrics.

A. Sleep period time window (SPT) validation

The GGIR processing algorithm uses a heuristic algorithm looking at distribution change in z-angle of the forearm to detect sleep onset time and wake time. Sleep onset time, waking time and sleep period time (SPT) window (time from sleep onset to waking) were validated in a 2018 paper [9] on 50 patients recorded with polysomnography and 3752 participants recorded with sleep logs. Sleep onset time had a mean absolute error (MAE) of 39.9 minutes, wake time had a MAE of 29.9 minutes and the SPT window had an MAE of 40 minutes compared to sleep log data. Validation results are shown in Table I compared to a previous algorithm that did not rely on the forearm angle for sleep detection.

TABLE I
MEAN ABSOLUTE ERROR OF SLEEP ONSET TIME, WAKE TIME AND SLEEP PERIOD TIME (SPT) WINDOW.

[h] **Algorithm**	**Sleep Onset** [min]	**Wake Time** [min]	**SPT Window** [min]
HDCZA with fore-arm angle	39.9	29.9	40
Heuristic without fore-arm angle	93.3	58.4	128.4

B. Physical activity level validation

A study found that the Verisense algorithm was accurate at detecting various physical activity levels in a population that consisted of thirty children (7-11 years) and thirty adults (18-65 years) [10]. Euclidean norm minus one values were calculated from the wrist sensor and a cut point approach was used to classify activities into sedentary, light physical activity (PA), moderate PA and vigorous PA. Testing was performed under supervised conditions on a treadmill. Classifications from the wrist sensor ENMO outputs were compared to metabolic equivalent of task (MET) values from the metabolic gas analysis system. Results are summarized in Table II. It was found that physical activity level could be accurately identified for all by slow walking conditions.

C. Wake after sleep onset

The sleep efficiency score is calculated as the ratio of nocturnal wake time in between sleep onset time and wake up time. This is calculated based on five minute windows of raw data from the tri-axial accelerometer. Nocturnal wake times were compared to polysomnography data from 28 patients and yielded an accuracy of 81%, a sensitivity of 81% and a specificity of 60% [11]. In this case, sensitivity represents the percentage of correctly identified sleep periods and specificity represents the percentage of correctly identified nocturnal wake periods. This research shows that sleep efficiency can be accurately obtained using GGIR processing.

TABLE II
VALIDATION RESULTS OF ACTIVITY LEVEL IDENTIFICATION FROM VERISENSE.

Physical Activity	**Detection Accuracy**
Laying	100%
Sitting	96%
Standing	100%
Slow Walk	55%
Fast Walk	100%
Running	97%

TABLE III

STUDIES PERFORMED USING GGIR DATA PROCESSING.

Therapeutic area	**Patient Cohort**	**Number of Patients (n)**
Cardiovascular	Heart surgery patients [44]	80
	Stroke [15]	41
	Cardiovascular Disease [16]	23,742
	Coronary Artery Disease [17]	58
Central Nervous System	Muscular Dystrophy [18]	128
	Dementia [19]	26
Musculoskeletal	Idiopathic inflammatory myopathy [20]	55
	Muscular Dystrophy [18]	128
	Sarcopenia [21]	131
Mental Health	Depression [22]	359
	Bipolar Disorder [23]	46
	Post-Partum Depression [24]	21
Diabetes	Gestational Diabetes [25]	697
	Type II Diabetes Mellitus [26]	635
	Type II Diabetes [27]	246
Rehabilitation and recovery	Pulmonary rehab patients [28]	79
	Bariatric surgery patients [29]	22
Pulmonary	Cystic Fibrosis [30]	9
	Idiopathic Pulmonary Fibrosis [31]	35
Aging	Older adults [32]	1,451
	Post-menopausal women [33]	1,316
Lifestyle	Sedentary adults [34]	191
	Obesity [35]	1,986
	Smoking [36]	3,063
	General population [37]	85,388
Children	Obese / overweight [38]	208
	Adolescents [39]	2,526
	Children [40]	2,636
Pregnant Women	Pregnant women [41]	2,317
	Pregnant and overweight [42]	257

IV. CLINICAL VALIDATION

The third and final stage of the V3 validation framework is clinical validation. Clinical validation evaluates whether a biometric monitoring technology acceptably identifies, measures or predicts a meaningful clinical, biological, physical, functional state or experience, in the stated context of use (which includes a specified population). In many cases, this context will include a specific symptom of interest as well (e.g. freezing versus tremors in Parkinson's patients). Because there are so many possible metrics, this validation may need to be performed in conjunction with the specific study to be undertaken.

One of the advantages of using a widely used open source software package like GGIR is that many studies (over 150 in the case of GGIR) have been performed on different populations and therapeutic areas. Table III summarizes the work performed in different therapeutic areas and cohorts. We highlight some specific results from this large body of evidence below.

A. Stroke

The relationship between physical activity, sleep and fatigue was assessed in forty one stroke patients using wrist worn accelerometry and the same processing algorithms that Verisense utilize [13]. Researchers found that stroke survivors performed less moderate to vigorous physical activity (MVPA) in ten minute bouts than the National Stroke guidelines recommend. They also found associations between light physical activity and fatigue as well as MVPA and sleep efficiency.

B. Cardiovascular Disease & Type II Diabetes

Researchers used a subset of data from the UK Biobank study (n = 106,053) to investigate the relationship between cardio-metabolic disease and physical activity in a large scale, objective study [14]. They found that men and women with the worst cardio-metabolic disease perform around half of moderate to vigorous physical activity on a daily basis compared to healthy individuals and spend almost 7 hours per day in 30 minute inactivity bouts. The researchers state that tri-axial accelerometers provide enhanced measurement opportunities for measuring lifestyle behaviors in chronic disease.

C. Dementia

Twenty six community dwelling people with mild dementia were asked to wear a wrist based monitor for thirty days for a feasibility and acceptance study [12]. Results indicated that patients tended to find wearing the activity monitors acceptable, with only three participants withdrawing prior to the end of the study. Dementia patients were satisfied with wearing the wrist device for one month as measured by the Quebec User Evaluation of Satisfaction with assistive Technology survey.

V. CONCLUSIONS

The Verisense IMU has undergone significant validation in all three levels of the V3 process. Although it is impossible for any device to be completely validated for every possible therapeutic use case, use of a commonly used open-source software package allows researchers to leverage previous validation work. Furthermore, any published results of new validation studies can be used to build up the body of validation work for future researchers - even if they do not use the same specific device. By using GGIR, the most widely used open source analysis package for wrist-based accelerometry, the Verisense validation effort can leverage an ever-growing body of validation work available.

REFERENCES

[1] Burns, Adrian, Barry R. Greene, Michael J. McGrath, Terrance J. O'Shea, Benjamin Kuris, Steven M. Ayer, Florin Stroiescu, and Victor Cionca. "SHIMMER™–A wireless sensor platform for noninvasive biomedical research." IEEE Sensors Journal 10, no. 9 (2010): 1527-1534.

[2] Shahzad, Ahsan, Seunguk Ko, Samgyu Lee, Jeong-A. Lee, and Kiseon Kim. "Quantitative assessment of balance impairment for fall-risk estimation using wearable triaxial accelerometer." IEEE Sensors Journal 17, no. 20 (2017): 6743-6751.

[3] Martinez-Manzanera, O., E. Roosma, M. Beudel, R. W. K. Borgemeester, Teus van Laar, and Natasha M. Maurits. "A method for automatic, objective and continuous scoring of bradykinesia." In 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN), pp. 1-5. IEEE, 2015.

[4] Lukšys, Donatas, Gintaras Jonaitis, and Julius Griškevičius. "Quantitative analysis of parkinsonian tremor in a clinical setting using inertial measurement units." Parkinson's Disease 2018 (2018).

[5] Billeci, L., G. Pioggia, E. Brunori, G. Crifaci, G. Tartarisco, R. Balocchi, S. Maestro, and M. A. Morales. "Wearable sensors combined with wireless technologies for the evaluation of heart rate and heart rate variability in anorexia nervosa adolescents." Neuropsychiatrie de l'enfance et de l'adolescence 5, no. 60 (2012): S157.

[6] Migueles, Jairo H., Alex V. Rowlands, Florian Huber, Séverine Sabia, and Vincent T. van Hees. "GGIR: a research community–driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data." Journal for the Measurement of Physical Behaviour 2, no. 3 (2019): 188-196.

[7] Goldsack, Jennifer C., Andrea Coravos, Jessie P. Bakker, Brinnae Bent, Ariel V. Dowling, Cheryl Fitzer-Attas, Alan Godfrey et al. "Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs)." npj digital Medicine 3, no. 1 (2020): 1-15.

[8] McDevitt, Bríd; Moore, Lisa; Akhtar, Nishat; Connolly, James; Doherty, Rónán; Scott, William. 2021. "Validity of a Novel Research-Grade Physical Activity and Sleep Monitor for Continuous Remote Patient Monitoring" Sensors 21, no. 6: 2034. https://doi.org/10.3390/s21062034

[9] van Hees, V.T., Sabia, S., Jones, S.E., Wood, A.R., Anderson, K.N., Kivimäki, M., Frayling, T.M., Pack, A.I., Bucan, M., Trenell, M.I. and Mazzotti, D.R., 2018. Estimating sleep parameters using an accelerometer without sleep diary. Scientific reports, 8(1), p.12975.

[10] Hildebrand, M.V.H.V., VAN, V.H., Hansen, B.H. and Ekelund, U.L.F., 2014. Age group comparability of raw accelerometer output from wrist-and hip-worn monitors. Medicine and science in sports and exercise, 46(9), pp.1816-1824.

[11] Van Hees, Vincent T., Séverine Sabia, Kirstie N. Anderson, Sarah J. Denton, James Oliver, Michael Catt, Jessica G. Abell, Mika Kivimäki, Michael I. Trenell, and Archana Singh-Manoux. "A novel, open access method to assess sleep duration using a wrist-worn accelerometer." PloS one 10, no. 11 (2015): e0142533.

[12] Farina, Nicolas, Gina Sherlock, Serena Thomas, Ruth G. Lowry, and Sube Banerjee. "Acceptability and feasibility of wearing activity monitors in community-dwelling older adults with dementia." International journal of geriatric psychiatry 34, no. 4 (2019): 617-624.

[13] Shepherd, Anthony I., Richard Pulsford, Leon Poltawski, Anne Forster, Rod S. Taylor, Anne Spencer, Laura Hollands et al. "Physical activity, sleep, and fatigue in community dwelling Stroke Survivors." Scientific reports 8, no. 1 (2018): 1-8.

[14] Cassidy, Sophie, Harley Fuller, Josephine Chau, Michael Catt, Adrian Bauman, and Michael I. Trenell. "Accelerometer-derived physical activity in those with cardio-metabolic disease compared to healthy adults: a UK Biobank study of 52,556 participants." Acta diabetologica 55, no. 9 (2018): 975-979.

[15] Shepherd AI, Pulsford R, et al. Physical activity, sleep, and fatigue in community dwelling Stroke Survivors. Scientific Reports 2018, doi: 10.1038/s41598-018-26279-7.

[16] Cassidy S, Fuller H et al Accelerometer-derived physical activity in those with cardio-metabolic disease compared to healthy adults: a UK Biobank study of 52,556 participants, Acta Diabetologica, 2018, doi: 10.1007/s00592-018-1161-8.

[17] Charman et al The effect of percutaneous coronary intervention on habitual physical activity in older patients. BMC Cardiovasc Disord. 2016 Dec 3;16(1):248.

[18] Okkersen K, Jimenez-Moreno C et al. Cognitive behavioural therapy with optional graded exercise therapy in patients with severe fatigue with myotonic dystrophy type 1: a multicentre, single-blind, randomised trial, Lancet Neurology, 2018, doi: 10.1016/S1474-4422(18)30203-5.

[19] Farina N, Sherlock G et al. Acceptability and feasibility of wearing activity monitors in community-dwelling older adults with dementia. Geriatric Psychiatry 2019. doi: 10.1002/gps.5064.

[20] Landon-Cardinal O, Bachasson D, et al. Relationship between change in physical activity and in clinical status in patients with idiopathic inflammatory myopathy: a prospective cohort study. Seminars in Arthritis and Rheumatism 2020. doi: 10.1016/j.semarthrit.2020.06.014.

[21] Westbury LD, Dodds RM et al. Associations Between Objectively Measured Physical Activity, Body Composition and Sarcopenia: Findings from the Hertfordshire Sarcopenia Study (HSS), Calcified tissue international, 2018, doi:10.1007/s00223-018-0413-5).

[22] Difrancesco S, Lamers F, et al. Sleep, circadian rhythm, and physical activity patterns in depressive and anxiety disorders: A 2-week ambulatory assessment study. Depression and Anxiety, 2019. doi: 10.1002/da.22949.

[23] Bradley AJ, Webb-Mitchell R, et al Sleep and circadian rhythm disturbance in bipolar disorder. Psychological Medicine 2017 Fenbruary, DOI: 10.1017/S0033291717000186.

[24] Warehime S, Dinkel D et al. Postpartum physical activity and sleep levels in overweight, obese and normal-weight mothers, British Journal of Midwifery 2018, doi: 10.12968/bjom.2018.26.6.400.

[25] Dawkins NP, et al. Comparing 24 h physical activity profiles: Office workers, women with a history of gestational diabetes and people with chronic disease condition(s) Journal of Sports Sciences 2020, doi: 10.1080/02640414.2020.1812202.

[26] Mickute M, et al. Device-measured physical activity and its association with physical function in adults with type 2 diabetes mellitus. Diabetic Medicine 2020. doi: 10.1111/dme.14393.

[27] Hausler N, Marques-Vidal P et al. Association between actigraphy-based sleep duration variability and cardiovascular risk factors – Results of a population-based study. Sleep Medicine. 2019. doi: 10.1016/j.sleep.2019.02.008.

[28] Chevance G, Berry Tanya, et al. Changing implicit attitudes for physical activity with associative learning. German Journal of Exercise and Sport Research, 2018. doi: 10.1007/s12662-018-0559-3.

[29] Afshar et al. Changes in physical activity after bariatric surgery: using objective and self-reported measures. Surg Obes Relat Dis. 2017 March. doi: 10.1016/j.soard.2016.09.012.

[30] Shelley J, Fairclough SJ, et al. A formative study exploring perceptions of physical activity and physical activity monitoring among children and young people with cystic fibrosis and health care professionals. BMC Pediatrics, 2018. doi: 10.1186/s12887-018-1301-x.

[31] Atkins C, Baxter M, et al. Measuring sedentary behaviors in patients with idiopathic pulmonary fibrosis using wrist-worn accelerometers. The Clinical Respiratory Journal 8 January 2017. DOI: 10.1111/crj.12589.

[32] Bielemann RM, LaCroix AZ, et al. Objectively Measured Physical Activity Reduces the Risk of Mortality among Brazilian Older Adults. J of the American Geriatrics Society 2019. doi: 10.1111/jgs.16180.

[33] Rowlands AV, Sherar LB, et al. A data-driven, meaningful, easy to interpret, standardised accelerometer outcome variable for global surveillance. Journal of Science and Med. in Sport 2019. doi: 10.1016/j.jsams.2019.06.016.

[34] Amaro-Gahete FJ, Acosta FM, et al. Association of sedentary and physical activity time with maximal fat oxidation during exercise in sedentary adults, Scan J of Med Science in Sports 2020 doi: 10.1111/sms.13696

[35] Papandreou C, Bullo M, et al. High sleep variability predicts a blunted weight loss response and short sleep duration a reduced decrease in waist circumference in the PREDIMED-Plus Trial. Int J of Obesity. 2019. doi: 10.1038/s41366-019-0401-5

[36] Benadjaoud MA, Menai M, et atl. The association between accelerometer-assessed physical activity and respiratory function in older adults differs between smokers and non-smokers. Nature Scientific Reports, 2019. doi: 10.1038/s41598-019-46771-y.

[37] Wang H, Lane JM, et al. Genome-wide association analysis of self-reported daytime sleepiness identifies 42 loci that suggest biological subtypes, Nature Communications 2019. doi: 10.1038/s41467-019-11456-7.

[38] Gomez-Bruton A, Arenaza L, et al. Associations of dietary energy density with body composition and cardiometabolic risk in children with overweight and obesity: role of energy density calculations, under-reporting energy intake and physical activity. British Journal of Nutrition. 2019. doi: 10.1017/S0007114519000278.

[39] Bielemann RM, Ramires VV, et al. Is vigorous-intensity physical activity required for improving bone mass in adolescence? Findings

from a Brazilian birth cohort. Osteoporosis International, 2019. doi: 10.1007/s00198-019-04862-6.

[40] Bielemann RM, dos S Vaz J et al. Are consumption of dairy products and physical activity independently related to bone mineral density of 6-year-old children? Longitudinal and cross-sectional analyses in a birth cohort from Brazil, Public Health Nutrition 2018, doi: 10.1017/S1368980018001258.

[41] Lloyd J, Creanor S, et al. Effectiveness of the Healthy Lifestyles Programme (HeLP) to prevent obesity in UK primary-school children: a cluster randomised controlled trial. BMC Public Health, 2017. DOI: 10.1186/s12889-017-4196-9.

[42] da Silva SG, Evenson KR, et al. Correlates of accelerometer-assessed physical activity in pregnancy:The 2015 Pelotas (Brazil) Birth Cohort Study. Scandinavian journal of medicine and science in sports, 2018. doi: 10.1111/sms.13083.

[43] Phelan S, Wing RR, et al. Randomized controlled clinical trial of behavioral lifestyle intervention with partial meal replacement to reduce excessive gestational weight gain. The American Journal of Clinical Nutrition, 2018, doi: 10.1093/ajcn/nqx043.

[44] Tillin T, Tuson C. et al. Yoga and Cardiovascular Health Trial (YACHT): a UK-based randomised mechanistic study of a yoga intervention plus usual care versus usual care alone following an acute coronary event. BMJ Open 2019. doi: 10.1136/bmjopen-2019-030119.

[45] Van Hees, Vincent T., Zhou Fang, Joss Langford, Felix Assah, Anwar Mohammad, Inacio CM da Silva, Michael I. Trenell, Tom White, Nicholas J. Wareham, and Søren Brage. ”Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents.” Journal of applied physiology 117, no. 7 (2014): 738-744.