Supplementary 6. Diagnostic Performance of Wearable Devices in Detecting Complications in Hematology and Oncology Patients

Clinical Complication	Definition	Wearable device	Wearable Device Measurement Used	Non-Wearable Device Measurement Used	Reference Standard	Artificial Intelligence Used	Accuracy	Article
SCC	CTCAE grade ≥3	Everion (wearable biosensor)	Heart rate, temperature, respiratory rate, oxygen saturation, blood pressure wave, and physical activity	NA	Clinical documentation of SCC	Self-supervised contrastive deep learning model	Inpatient cohort: AUROC: 0.91±0.01; sensitivity: 79.7%; specificity: 87.9% Outpatient cohort: AUROC: 0.87±0.02; sensitivity: 77.4%; specificity: 81.8%	6
Post-operative events	Pain, dehydration, constipation, and wound concerns	Fitbit Alta HR (wearable biosensor)	Lightly active minutes, fairly active minutes, sedentary minutes, mean daily heart rate, daily steps, and sleep	Patient-reported-outcomes questionnaire: physical function, sleep disturbance, anxiety, fatigue, and pain intensity	Actual occurrence of adverse events	Machine learning (predictive modelling)	AUC: 0.75 (95% CI: 0.67–0.81)	1
	Complications with Clavien-Dindo Grade 2 or higher were considered clinically relevant deteriorations	HealthDot (adhesive patch)	Heart rate and respiratory rate	NIL	Conventional MEWS derived from patient monitor in the intensive and post-care unit	NIL	AUC: 0.71 (95% CI: 0.66–0.77)	16
							Clarke error grid analysis showed that 100% of the hazard ratio, and 99.4% of the 5 minutes averaged data were clinically acceptable	17
CRS	A severe systemic inflammatory reaction that can occur in	Current Health Inc.	Temperature, pulse, respiratory rate, and	NIL	Visual inspection and inter-observer	NIL	Detect temperature change at a median of	10

	patients receiving CAR-T cell therapy or HCT	(wearable biosensor)	oxygen saturation		correlation		205 minutes earlier	
Infection/ Sepsis	A temperature ≥ 38 °C with no fever recorded in the past 24 hours at least 3 recordings ≥ 38°C within the first hour of fever, and that the fever duration must be greater than 1 hour Lee et al., (2019)	TempTraq (adhesive patch)	Temperature	NIL	Standard-of-care intermittent temperature monitoring	NIL	Detected 89% of fevers in a median of 5.5 hours earlier	4
Neutropenia	Confirmed infection: positive microbiologic culture result or positive radiographic imaging for infection Febrile neutropenia: temperature >38.3 °C measured once or >38.0 °C measured twice over 1 hour in the setting of a neutrophil count <500 cells/mL.	Zephyr BioPatch/ BioHarness (wearable biosensor)	Heart rate variability	Laboratory biomarker data: CRP, PCT, sCD163, IL-7, soluble TREM-1, IL-1b, IL-6, IL-8, and TNFα.	The adjudication of antibiotic escalation by two blinded HCT clinicians	Logistic regression model and combined rule-based naïve Bayes model	(Without laboratory biomarker data): AUROC: 0.66 Specificity: 0.88 Sensitivity: 0.28 (With laboratory biomarker data): AUROC: 0.87 Specificity: 0.86 Sensitivity: 0.68	2
	NIL	TempTraq (adhesive patch)	Temperature	NIL	Standard-of-care intermittent temperature monitoring	NIL	AUC of binary temperature skin patch in all four-time intervals was significantly higher than SOC: 0–30 mins AUC: 0.766 (CI: 0.708–0.824, P<0.001);	3

							30mins–1 hour AUC: 0.755 (CI: 0.701–0.809, P<0.001); 1–2-hour AUC: 0.718 (CI: 0.663–0.773, P<0.001); 2–4-hour AUC: 0.702 (CI: 0.646–0.757, P<0.001)	
Death events	End of life	Garmin VivoSmart 4 (wearable biosensor)	Heart rate, sleep, oxygen saturation, steps, and sleep	Consciousness, appetite, urination, edema, pain score, sleep, drowsiness, nausea, constipation, diarrhea, dyspnea, fatigue, fever, functional level (using Australia-modified Karnofsky Performance Status), care phase, pain control change, and basic demographic	The actual occurrence of death within the next 7 days	XG Boost Model	AUROC: 96%, F1-score: 78.5%, Accuracy: 93%, Specificity: 97%"	8
Pain	Daily chronic pain as well as intermittent, unpredictable acute vaso-occlusive painful episodes called pain crises	Microsoft Band 2 (Smartwatch)	Heart rate, galvanic skin response, and angular velocity	TRU-Pain app, which allows patients to record pain, general health and mood using visual analog scale.	Pain score reported by patients	Support vector machines for regression	SVM model Accuracy: 0.682 F1 score of mild pain: 0 F1 score of moderate pain: 0.537 F1 score of severe pain: 0.786 Weighted F1 score: 0.663	7

							SVM for regression model Accuracy: 0.729 F1 score of mild pain: 0.286 F1 score of moderate pain: 0.675 F1 score of severe pain: 0.803 Weighted F1 score: 0.728	
	Episodes of acute and severe pain known as vaso-occlusive crises	Apple Watch (Smartwatch)	Heart rate, heart rate variability, oxygen saturation, and activity levels	Self-reported pain scores using visual analog scale collected via the Nanpar mobile app. Pain scores were recorded on a visual analog scale ranging from 0 to 10, with 0 accounting for no pain, and 10 being most intense pain.	Pain scores from the electronic health record	Machine learning Random Forest classification model	Micro-averaged accuracy: 0.89, (89%) Micro-averaged F1-score: 0.49 RMSE: 1.64 AUC: 0.83	18
	Episodes of acute and severe pain known as vaso-occlusive crises	Apple Watch (Smartwatch)	Heart rate, heart rate variability, calories	Pain scores and vital sign variables including blood pressure, pulse, and temperature, as well as demographics including age, SCD genotype, sex, and ethnicity were	Nurse-recorded pain score from electronic health record	Multinomial logistic regression, gradient boosting, and random forest machine learning models	Random forest model: AUC: 0.98 Accuracy: 84.5% Micro-averaged F1 score: 0.85 RMSE: 0.84	13

				extracted from the electronic medical records				
Fatigue	Diminished energy and mental capacity, and disturbed psychological conditions among patients with cancer	ViPCare, Gadgletech (wearable biosensor)	Heart rate variability	BFI- Taiwan version questionnaire	BFI- Taiwan version questionnaire	NIL	The total mapping error rate was 3% and the LF-HF ratio can be considered a fair indicator to evaluate the degree of cancer-related fatigue during cancer treatment	11

AUC: area under curve; AUROC: area under the receiver operating characteristic; BFI: brief fatigue inventory; CAR-T: chimeric antigen receptor T-cell; CRP: C-reactive protein; CRS: cytokine release syndrome; CTCAE: common terminology criteria for adverse events; HCT: hematopoietic cell transplantation; PCT: procalcitonin; LF-HF : low-frequency to high-frequency; MEWS: modified early warning score; RSME: root mean square error; SCC: serious clinical complications; SCD: sickle cell disease; sCD163: soluble CD163; SOC: sTREM-1: soluble TREM-1; SVM: support vector machine; TRU-Pain: technology resources to understand pain.

Apple, Cupertino, California, USA; Current health inc, Edinburgh, UK; Everion, Biovotin AG, Zürich, Switzerland; Fitbit, San Francisco, California, USA; Garmin: Olathe, Kansas, USA; HealthDot, Philips, Amsterdam, the Netherlands; Microsoft: Redmond, Washington, USA; TempTraq: Westlake, Ohio, USA; ViPCare, Gadgletech, London, UK; s: Zephyr™: Boulder, Colorado, USA