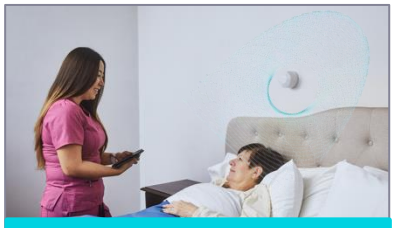
XANDAR KARDIAN

AI-DRIVEN HEALTH RISK DETECTION

Large-scale dataset of real health outcomes

XK's sensors continuously and passively capture heart rate, respiratory rate, movement, sleep and more. These long-term vital sign trends have been linked to health outcomes recorded in clinical care settings. Together, this data supports the development and evaluation of AI models for risk detection.



The flagship XK300 Sensor

5.2
Billions of heart rate captured

9.6
Billions of respiratory rate captured

2500
Patient Years of total monitoring

7000
Health outcomes observed

Know before you know

The sensor can detect subtle physiological changes that signal the approach of a new health condition, even while vital sign readings still appear within normal ranges. These patterns are learned directly from sensor data and linked to real outcomes, allowing AI models to estimate risk over time. This allows earlier visibility into emerging issues and additional time to monitor, assess, or intervene. With Xandar Kardian, know before you know.

Health Outcome	Cases Seen (monitored by XK's sensors*)	Model Accuracy (when specificity = sensitivity**)	Early Notice (averaging first alerts per case)
Hospitalization	510	76.5 %	5.5 days
Death	335	89.5 %	6.3 days
Covid-19	189	71.1 %	3.9 days
UTI	108	74.5 %	4.3 days
COPD	31	72.5 %	4.1 days
Heart Failure	28	73.2 %	5.3 days

*Cases were included only if patients had at least 5 out of the last 7 days with sufficient presence under the sensor (>1000 HR/RR, or about 3 hours of monitoring per day).

**Accuracy was measured on never-seen before data, at the threshold where sensitivity equals specificity, ensuring prevalence-independent reliability; the model can later be tuned to favor either metric.

www.xk.inc
sales@xk.inc
5000 Yonge St, Suite 1001
Toronto, Canada

SCIENTIFIC METHODOLOGY

1. Study Context and Data Sources

1.1 Care Setting Demographics

All data used in this analysis were collected retrospectively in U.S.-based long-term care (LTC) facilities. The results presented reflect performance within this specific population and care environment. In total, these studies leveraged the data of 2,402 patients across 34 facilities, with ages averaging 78 years old. The data collected by the sensors spanned from 2022 to 2025 and totaled over 22,300 days of monitoring.

1.2 Physiological Data Collection

XK sensors passively collect continuous physiological measurements, such as:

- Heart rate
- Respiratory rate
- Movement
- Sleep-related metrics

The sensor measures heart rate and respiratory rate every three seconds. A built-in signal validation process evaluates each measurement for stability, and only readings meeting predefined quality criteria are retained to guarantee accuracy.

This filtering process is designed to reduce the impact of disturbances caused by other sources of movement, such as non-patient motion, temporary signal obstruction, or environmental vibrations.

The sensor is typically installed above the bed, where the patient is expected to be the primary and closest source of motion. Obstructions placed directly between the patient and the sensor, particularly reflective materials such as metal or glass, may degrade signal quality, while blankets and clothing do not interfere with signal acquisition.

Minor hardware variations may exist across device revisions; however, firmware and signal-processing logic are standardized, and devices are validated to ensure equivalent performance.

1.3 Health Outcome Identification

Health outcomes are derived from electronic health record (EHR) systems used within participating LTC facilities.

Structured EHR data includes:

- Admission, discharge, and transfer (ADT) events
- Primary diagnoses (e.g., UTI, COVID-19, CHF)
- Medication records
- Care documentation

Hospital transfers and deaths are identified through standardized discharge fields (ADT), with predefined categorical values.

Diagnoses are recorded as structured primary diagnosis entries by licensed care staff, including nurses and physicians. They are documented using standardized ICD-10 classification codes, ensuring consistency across clinical staff and facilities.

The timestamps attached to each outcome are recorded within the EHR system. For discharge events, timestamps are typically precise. For diagnoses, timestamps reflect the time of clinical documentation and may occur after initial symptom onset.

All data are anonymized prior to model development.

2. Dataset Construction

2.1 Defining Study and Control Groups

For each health outcome, a study group is formed from individuals with a documented occurrence of that outcome. Sensor data from a defined time window preceding the documented event are extracted.

The duration of the pre-event window varies by outcome type and is selected based on expected symptoms onset and empirical clinical observation:

- Shorter windows (e.g., 5 days) for acute infectious conditions
- Longer windows (e.g., 7–14 days) for outcomes associated with gradual decline

For comparison, a control group of equal size is constructed, composed of individuals who did not experience the same outcome within a comparable time frame. Sensor data from an equivalent time window are extracted using identical selection and preprocessing procedures, ensuring methodological consistency between groups.

This produces a balanced dataset containing equal numbers of outcome and non-outcome examples. Although this balanced structure does not reflect real-world prevalence, it still enables unbiased evaluation of detection performance.

2.2 Handling Overlapping or Ambiguous Cases

- Cases with overlapping or confounding primary diagnoses may be excluded depending on outcome definitions.
- Suspected but undocumented outcomes cannot be independently verified and are therefore classified according to available records. As a result, some individuals in the control group may have experienced undocumented outcomes. This introduces potential label uncertainty and may lead to conservative performance estimates.
- Medication changes are not explicitly modeled as separate variables, physiological changes caused by such events are captured and can influence measurements.

3. Model Development and Evaluation

3.1 Training Procedure

Models are trained and evaluated using cross-validation splits. For each split:

- 80% of data is used for model training.
- The remaining 20% is reserved for evaluation.
- No patient can appear in both training and evaluation datasets.

The experiment is repeated multiple times with different random splits to confirm consistency of results and reduce variance.

Models are trained exclusively on features derived from sensor data. No demographic variables or external clinical inputs are used. This design enables deployment in environments where structured EHR integration may be limited or inconsistent.

3.2 Performance Metric

Model performance is evaluated by measuring:

- Sensitivity: the proportion of outcome cases correctly identified within the study group.
- Specificity: the proportion of non-outcome cases correctly identified within the control group.

The model produces a continuous risk score rather than a fixed yes/no decision. A decision threshold is applied to this score to determine when an alert is generated.

Adjusting this threshold changes the balance between sensitivity and specificity. Lowering the threshold increases detection accuracy (higher sensitivity) but increase false alerts (lower specificity). Raising the threshold reduces false alerts but decrease true detections.

For reporting purposes, accuracy is calculated at the operating point where sensitivity and specificity are equal. This provides a neutral reference point in which detection performance and false-alert control are given equal weight.

Because the evaluation dataset is balanced and performance is assessed at equal sensitivity and specificity, the reported accuracy reflects the model's intrinsic discrimination capability rather than the underlying prevalence of the outcome. In practical deployment, the threshold may be adjusted to align with operational priorities, such as emphasizing earlier detection or minimizing alert burden.

3.3. Estimation of Early Detection Window

Each day before a documented outcome is evaluated separately. For each day, the model uses that day's physiological measurements together with recent trends to produce a risk score.

If the risk score exceeds the alert threshold for a patient who later experiences the outcome, that day is considered a correct early detection. The first day on which this occurs is recorded.

Reported lead time represents the average number of days between this first correct detection and the documented outcome across cases.

4. Study Boundaries and Considerations

Results are based on retrospective analysis, and all data originate from U.S. long-term care setting. Performance may vary in other populations or care settings.

Documentation timing in EHR systems may not coincide exactly with physiological onset. As individual physiological responses vary, model performance should be monitored over time, and calibration may be required as care setting evolve.

The sensor and associated models are intended to support clinical monitoring by identifying patterns that may warrant closer review. They are not intended to independently establish medical diagnoses or initiate clinical decisions without human evaluation. Alerts should be interpreted within the context of facility workflows and clinical judgment.