Clinical Brief

Understanding Sleep Diagnostics with DeepResp by Nox Medical

LBL-0698 REV01

Understanding Sleep Diagnostics with DeepResp by Nox Medical

AI-Assisted Sleep Staging and Arousal Scoring for More Conclusive HSAT

Precise measurement of breathing is fundamental to sleep-diagnostics.[1,2] Beyond identifying reductions in airflow that define apneas and hypopneas, breathing patterns also reflect underlying sleep physiology. Sleep stages and arousals produce characteristic changes in ventilation and respiratory effort, creating physiological signatures that can be detected from respiratory signals. Accurately capturing these changes is therefore essential not only for identifying respiratory events, but also for estimating sleep state and arousals from breathing itself.

For more than 20 years, Nox Medical has focused on advancing the measurement of breathing as the foundation of its sleep diagnostics technology. Through continuous innovation in respiratory inductance plethysmography (RIP), Nox has developed a volume-based measurement technology (volumetric RIP)[3] and calibrated derived signal (Nox Flow)[5] – transforming thoracoabdominal signals from dual RIP belts into a reliable[4], physiologically grounded estimate of airflow, closely aligned with gold-standard pneumotachography.[5] This approach enables reliable quantification of ventilation and respiratory effort with high accuracy — precision that supports analysis beyond conventional HSAT.[3]

This document explains how Nox DeepResp v.2.0 technology works, including its scientific validation and the clinical value it delivers. Using DeepResp to score HSAT (including scoring sleep states and arousals, in addition to respiratory events), provides an AHI closely aligned with PSG.[23] DeepResp therefore supports a reduction in inconclusive or false-negative HSAT results.

The Physiology of Breathing During Sleep

Obstructive sleep apnea (OSA) is, at its core, a disorder of airflow. Clinical definitions of apnea and hypopnea are based on reductions in airflow associated with either oxygen desaturation or arousal — making both precise respiratory measurement and knowledge of sleep architecture essential for accurate diagnosis. Crucially, sleep stages and arousals can themselves be estimated from breathing patterns: subtle but distinctive changes in ventilation and respiratory effort closely correlate with wakefulness, sleep stages, and arousals, providing a robust physiological signal source for estimating sleep architecture directly from respiratory signals alone:

Wake[12,13]	Breathing is typically irregular. The upper airway is generally patent, with neither snoring nor respiratory events present.
NREM Sleep[12,14]	Breathing is regular and metabolically driven. Upper airway patency is reduced, increasing upper airway resistance and the consequent likelihood of snoring and respiratory events.
REM Sleep[12,14]	Breathing is driven primarily by the diaphragm due to atonia of the accessory muscles of respiration, and is more variable due to autonomic instability. Upper airway patency is further reduced, increasing the likelihood of respiratory events.
Arousals[12,13]	These events trigger characteristic respiratory responses, including muscle activation with brief surges in ventilation; both tidal volume and respiratory rate are increased.

Limitations of Traditional HSAT

Traditional home sleep apnea tests (HSAT) do not include EEG and therefore cannot directly measure sleep stages or cortical arousals. This creates two important limitations:

- Respiratory indices are typically calculated using total recording time or estimated time in bed rather than true sleep time, which can lead to an underestimation of disease severity.
- HSAT scoring has traditionally relied on oxygen desaturation alone to identify hypopneas. This means arousal-based hypopneas go undetected. These are events with flow limitation significant enough to cause an arousal but without scorable desaturation.[15]

These limitations affect most cases of OSA but are most clinically significant in populations whose disease is characterized by subtle airflow limitation and arousal-based physiology: women, younger adults, non-obese individuals, patients with REM-related OSA, and those with comorbid insomnia (COMISA).[16,17] For these patients, traditional HSAT is likely to yield false-negative or inconclusive results, delaying diagnosis and appropriate treatment. The AASM Hypopnea Scoring Rule Task Force has highlighted that inclusion of arousal-based hypopnea scoring affects OSA prevalence, treatment eligibility, and identification of patients who may otherwise remain undiagnosed.[18] Standard HSAT systems can only apply the desaturation arm of the AASM recommended hypopnea rule, meaning arousal-terminated events go undetected.

Why is it important to score arousals?

Arousals are:

- Brief several-second interruptions of sleep that are not consciously experienced as awakenings
- Triggered by physiologic (e.g., breathing obstruction) or environmental (e.g., auditory, tactile) stimuli

Consequences of arousals:

- Sleep fragmentation, leading to associated daytime impairment
- Activation of the sympathetic nervous system ("fight-or-flight" response) causing increases in:
  - Heart rate
  - Blood pressure
  - Serum glucose

Populations often underdiagnosed by traditional HSAT, such as women, younger patients, non-obese individuals, patients with co-morbid conditions, and those with REM-related OSA, frequently experience events with arousals rather than desaturation events. By scoring arousals, it is possible to capture clinically meaningful breathing disturbances that standard HSAT will miss.

What Is DeepResp?

Nox's FDA-cleared and CE-marked software-as-a-medical-device (SaMD), DeepResp™, is a cloud-based AI analysis platform that is validated for and supports both polysomnography (PSG) and home sleep apnea testing (HSAT) workflows. When EEG and EOG signals are available, as in PSG, DeepResp performs automated sleep staging and arousal scoring using standard neurophysiological signals. In HSAT, where EEG is not available, DeepResp uses Nox BodySleep v.2.0 technology to estimate sleep states and arousals from respiratory signals derived from Nox RIP technology to generate key clinical indices, complemented by advanced informational parameters.

This document focuses on the HSAT workflow within DeepResp, integrated into the Nox Connect workflow. Once a compatible Nox T3 or Nox A1 study is uploaded to Nox Connect, the study is automatically analyzed and scored by DeepResp and made available for clinician review in Noxturnal Web. Clinicians can review, edit, validate, and finalize scoring.

How DeepResp Works for HSAT scoring:

Sleep Staging and Arousal Scoring from RIP Alone

Using only Nox RIP signals, DeepResp enables sleep stage estimation and arousal scoring within HSAT, adding clinical information that standard HSAT cannot provide. Its scientific basis is rooted in physiology: by capturing volumetric airflow and thoracoabdominal movement from dual RIP belts, RIP technology provides quantitative signals related to effort and ventilation[3], from which sleep states and arousals can be scored.

For HSAT studies, the platform provides a preprocessing protocol, including:

Position detection
Identifies patient positions (upright, prone, supine, left and right positions).

Analysis period finder
Determines the likely period the patient attempted to sleep.

Desaturation detector
Identifies significant drops in blood oxygen saturation.

Nox BodySleep™ v.2.0
Advanced AI model using deep learning to analyze Nox RIP data, enabling scoring of sleep states (Wake, NREM, REM) and arousal events from respiratory signals without EEG. Supports more conclusive HSAT interpretation by determining actual sleep time and identifying hypopneas resulting in arousal without desaturation.

Nox Respiratory Analysis
Detects and scores various AASM-defined respiratory events, including obstructive and central apneas and hypopneas.

The HSAT postprocessing protocol includes:

Parameter engine
Computes sleep parameters, such as AHI.

Sleep apnea endotypes
Predicts the underlying causes of sleep apnea.
Informational purposes only. Not intended for clinical decision-making.

Sleep apnea burden
Ventilatory burden
Calculates the burden of airflow reduction by capturing the length and depth of respiratory events in one metric for a more detailed view of disease severity. Ventilatory burden has been shown in published research to predict cardiovascular and all-cause mortality.[24-26]
Informational purposes only. Not intended for clinical decision-making.

Sleep apnea burden
Hypoxic burden
Combines the depth and length of oxygen desaturation events to reflect the impact of breathing disruptions. Hypoxic burden has been shown in published research to predict cardiovascular morbidity and mortality.[6-11]
Informational purposes only. Not intended for clinical decision-making.

DeepResp outputs annotations for sleep stages, arousals and respiratory events, which the clinician can review together with the AHI and other relevant clinical indices. Additionally, DeepResp leverages ventilation signals from the nasal cannula or RIP flow, together with sleep and event analysis, to estimate key OSA endotypic traits – including loop gain, airway collapsibility, muscle compensation, and arousal threshold as well as burden parameters. These AI tools are provided for informational purposes only; they are not intended for clinical decision-making and are not supported by clinical claims.

DeepResp, including Nox BodySleep™ v.2.0, uses a deep learning model to analyze these breathing patterns and:

- Score 30-second epochs as Wake, NREM, or REM
- Detect and score arousals on a second-by-second basis
- Estimate total sleep time (TST) for use in AHI calculation
- Identify respiratory events, including hypopneas associated with arousals, enabling application of the full AASM recommended hypopnea rule

This allows HSAT to capture events traditionally missed when relying only on oxygen desaturation, and enables calculation of an AHI that closely aligns with expert-scored PSG.[1,5]

AASM Alignment: AASM scoring rules for hypopneas were developed in the context of in-lab PSG with EEG. Advances in technology and AI now enable arousal scoring in HSAT using AASM-recognized respiratory signals such as RIP.

In the context of HSAT with the Nox T3, FDA-cleared DeepResp uses analysis of Nox RIP technology to score arousals and sleep stages, extending arousal-based hypopnea identification to a broader clinical workflow, without EEG.

DeepResp uses traditional sleep study signals as inputs as recommended by the AASM. But sleep staging and arousal scoring is derived from secondary signals based on respiratory physiology (Nox RIP technology), conceptually familiar to healthcare professionals trained in AASM manual scoring.

Validation and Performance

All DeepResp validation compares automated scoring against manually scored PSG by RPSGT-certified sleep technicians according to AASM criteria. DeepResp is one of the most rigorously validated AI tools in sleep diagnostics, tested against a large and diverse dataset of sleep recordings from U.S. sleep clinics collected during routine clinical care. The FDA's Center for Devices and Radiological Health (CDRH) highly encourages the use of real-world evidence (RWE) in medical device applications, as it provides a deeper and more realistic understanding of the device's safety and effectiveness than traditional clinical studies alone. They recently published a report in which DeepResp was named as an example of a good use of RWE.[27]

The validation dataset for DeepResp v.2.0 included over 4,000 Type I in-lab PSG recordings and approximately 1,700 Type II recordings (HSAT with an EEG channel added to enable ground-truth comparison), spanning adult patients aged 18 and older across a wide range of ages (18-21, 22-35, 36-45, 46-55, 56-65, and 65+), body mass index values (<25, 25-29, >=30), and diverse geographic and racial/ethnic backgrounds, representative of a typical clinical population.

DeepResp was evaluated across two HSAT configurations reflecting common real-world setups: one including both nasal cannula airflow signals and Nox RIP belts (HSAT with cannula), and one based on Nox RIP signals from belts only (HSAT without cannula), to support obtaining a scored study even when the cannula signal is unavailable. Both configurations require an oximetry signal.

Performance

DeepResp is a clinically validated tool that supports confident, accurate decision-making. It improves true case detection, reducing the risk of underdiagnosis while delivering reliable results that support consistent patient care.

DeepResp's HSAT performance shows a high true-positive detection across all degrees of OSA severity (91–93.7% sensitivity for mild; 71.8–81% for moderate to severe) while maintaining strong specificity (up to 91.1%), helping rule out non-OSA cases and reduce unnecessary follow-up.

DeepResp v.2.0's arousal scoring shows 64% sensitivity, 90.5% specificity, and 83.1% accuracy. DeepResp reflects the known variability of arousal scoring, even among trained professionals.[19-22]

While positive agreement for arousal scoring may appear modest, this reflects a well-established characteristic of arousal scoring itself: published multi-scorer studies demonstrate that even RPSGT-certified sleep technologists show moderate agreement, with no single "ground truth."

Clinically Validated Performance

DeepResp automated scoring, including AI sleep staging and arousal detection alongside respiratory event scoring, demonstrated high diagnostic accuracy across clinical AHI thresholds. Performance metrics from a subset of the validation dataset indicate that DeepResp may identify more patients with OSA than traditional HSAT scoring based on rule-based respiratory event index (REI) alone.

AHI/REI ≥ 5 (Mild to severe OSA)	With cannula	Without cannula
	DeepResp (v.2.0) AHI (5,771 recordings)	
PPA Sensitivity (True Positive Rate)	91%	93.7%
NPA Specificity (True Negative Rate)	78%	63.5%

AHI/REI ≥ 15 (Moderate to Severe OSA)	With cannula	Without cannula
	DeepResp (v.2.0) (5,771 recordings)	
PPA Sensitivity (True Positive Rate)	71.8%	81%
NPA Specificity (True Negative Rate)	93.9%	91.1%

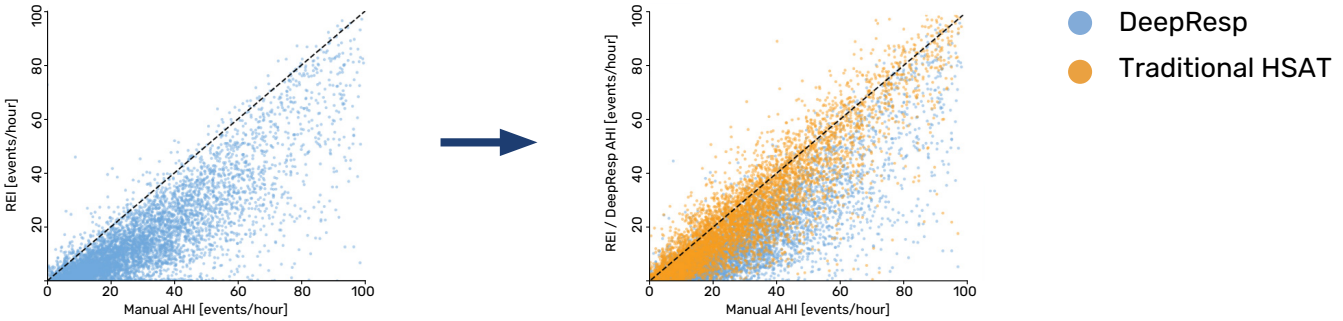

S

Clinical Value and Benefits

AHI Aligned with PSG

Scoring sleep and wake periods allows wake time to be excluded from the AHI calculation, producing a metric based on true sleep time rather than recording time. Combined with the ability to identify hypopneas associated with arousals, this enables calculation of an AHI that closely aligns with expert-scored PSG, markedly reducing inconclusive or false-negative HSAT results.

To illustrate, in a cohort of 5,771 real-world patients[23], traditional HSAT correctly identified OSA severity (mild to severe) in 77% of cases, missing nearly 1 in 4 patients. DeepResp improves diagnostic alignment with PSG to 92%, leading to a reduced number of missed or underestimated cases. In a cohort of 5,771, this translates to 827 additional patients correctly classified using DeepResp compared to traditional HSAT.



Differentiation of Apnea Types

Obstructive apneas and central apneas are recognized as distinct event types, aligning automated scoring with AASM diagnostic criteria. The Central Apnea-Hypopnea Index (CAHI) is provided alongside AHI, giving clinicians the specificity they expect when reviewing a scored study. Differentiating central sleep apnea from OSA is essential for selecting appropriate treatment.

Earlier Treatment Decisions

More accurate AHI reduces inconclusive and false-negative HSAT results, reducing the need for unnecessary follow-up PSG testing and enabling earlier, more confident diagnostic and treatment decisions.

Closing Diagnostic Gaps

DeepResp helps close diagnostic gaps within populations that are often underdiagnosed by traditional HSAT, including women, younger patients and non-obese individuals. These groups frequently experience respiratory events associated with arousals rather than desaturations. By scoring arousals, DeepResp can capture clinically meaningful breathing disturbances that standard HSAT cannot detect. DeepResp performance is robust across diverse populations; the groups mentioned above benefit particularly because AI sleep staging and arousal scoring address the specific limitations of traditional HSAT.

Transparent, Clinician-Led AI

DeepResp uses traditional sleep study signals as inputs and outputs parameters familiar to healthcare professionals trained in AASM manual scoring. Clinicians have full control: they can review underlying signal data, view probability and confidence levels for each scored sleep stage, arousal, or respiratory event, and manually adjust or override scoring as needed prior to diagnosis.

The model does not self-update or continuously learn after release, ensuring consistent and validated performance. All AI-generated results are intended to support, not replace, clinical judgment and must be reviewed by a qualified healthcare professional before making a diagnosis.

More Conclusive HSAT with DeepResp

By expanding what can be measured from respiratory signals, DeepResp brings greater diagnostic precision to home sleep testing. With DeepResp, clinicians can diagnose sleep apnea from a HSAT with greater confidence, particularly in patients often underdiagnosed by traditional approaches. By distinguishing between Wake, REM, and NREM, scoring arousals, and classifying obstructive and central events, clinicians also gain a more comprehensive understanding of a patient's sleep for better patient management. DeepResp helps prevent misclassification of severity leading to underdiagnosis, and identifies the full extent of OSA through robust respiratory event detection and sleep stage estimation, providing greater confidence in clinical decision-making while reducing delays in care and unnecessary retesting.

References

1. Finnsson E et al. Detecting arousals and sleep from respiratory inductance plethysmography. Sleep Breath. 2025 Apr 11;29(2):155.
2. Troester MM et al. The AASM Manual for the Scoring of Sleep and Associated Events (latest edition). American Academy of Sleep Medicine.
3. Bjarkason, S., Finnsson, E., Sigþórsson, S. et al. Toward Physiology-based Sleep Assessment: the Emerging Clinical Role of Advanced RIP Technology. Curr Pulmonol Rep 15, 10 (2026). https://doi.org/10.1007/s13665-026-00410-w
4. Montazeri K, Jonsson SA, Agustsson JS, Serwatko M, Gislason T, Arnardottir ES. The design of RIP belts impacts the reliability and quality of the measured respiratory signals. Sleep Breath. 2021 Sep;25(3):1535-1541. doi: 10.1007/s11325-020-02268-x. Epub 2021 Jan 7. PMID: 33411184; PMCID: PMC8376735.
5. Finnsson E et al. Respiratory inductance plethysmography to quantify changes in ventilation in obstructive sleep apnea. IEEE Trans Biomed Eng. 2025.
6. Peker Y, Celik Y, Zinchuk A, et al. Association of Hypoxic Burden With Cardiovascular Events. CHEST, 2025; 168, 1481-1493.
7. Azarbarzin A, Sands SA, Stone KL, Taranto-Montemurro L, Messineo L, Terrill PI, Ancoli-Israel S, Ensrud K, Purcell S, White DP, Redline S, Wellman A. The hypoxic burden of sleep apnoea predicts cardiovascular disease-related mortality: the Osteoporotic Fractures in Men Study and the Sleep Heart Health Study. Eur Heart J. 2019 Apr 7;40(14):1149-1157. doi: 10.1093/eurheartj/ehy624. Erratum in: Eur Heart J. 2019 Apr 7;40(14):1157. doi: 10.1093/eurheartj/ehz028. PMID: 30376054; PMCID: PMC6451769.
8. Blanchard M, Gervès-Pinquié C, Feuilloy M, Le Vaillant M, Trzepizur W, Meslier N, Goupil F, Pigeanne T, Balusson F, Oger E, Sabil A, Girault JM, Gagnadoux F; ERMES study group. Hypoxic burden and heart rate variability predict stroke incidence in sleep apnoea. Eur Respir J. 2021 Mar 25;57(3):2004022. doi: 10.1183/13993003.04022-2020. PMID: 33214210.
9. Azarbarzin A, Sands SA, Taranto-Montemurro L, Vena D, Sofer T, Kim SW, Stone KL, White DP, Wellman A, Redline S. The Sleep Apnea-Specific Hypoxic Burden Predicts Incident Heart Failure. Chest. 2020 Aug;158(2):739-750. doi: 10.1016/j.chest.2020.03.053. Epub 2020 Apr 13. PMID: 32298733; PMCID: PMC7417383.
10. Kim JS, Azarbarzin A, Wang R, et alAssociation of novel measures of sleep disturbances with blood pressure: the Multi-Ethnic Study of AtherosclerosisThorax 2020;75:57-63.

11 Jackson CL, Umesi C, Gaston SA, et alMultiple, objectively measured sleep dimensions including hypoxic burden and chronic kidney disease: findings from the Multi-Ethnic Study of AtherosclerosisThorax 2021;76:704-713.
12 Malik V, Smith D, Lee-Chiong T. Respiratory Physiology During Sleep. Sleep Medicine Clinics, 7, 497-505
13 Amatoury J, Jordan AS, Toson B, Nguyen C, Wellman A, Eckert DJ. New insights into the timing and potential mechanisms of respiratory-induced cortical arousals in obstructive sleep apnea. Sleep. 2018 Nov 1;41(11):zsy160. doi: 10.1093/sleep/zsy160. PMID: 30137568; PMCID: PMC6231527.
14 Tabachnik E, Muller NL, Bryan AC, Levison H. Changes in ventilation and chest wall mechanics during sleep in normal adolescents. J Appl Physiol Respir Environ Exerc Physiol. 1981 Sep;51(3):557-64. doi: 10.1152/jappl.1981.51.3.557. PMID: 7327955.
15 Scheuller HS et al. "Modified scoring criteria to improve the accuracy of the home sleep apnea test." Sleep and Breathing 30(1):52, 2026. DOI: 10.1007/s11325-026-03598-y
16 Malhotra RK. Arousal-based scoring of obstructive hypopneas. Curr Opin Pulm Med. 2021 Nov 1;27(6):491-495. doi: 10.1097/MCP.0000000000000820. PMID: 34410227.
17 Pitkänen M, Nath RK, Korkalainen H, Nikkonen S, Mahamid A, Oksenberg A, Duce B, Töyräs J, Kainulainen S, Leppänen T. Respiratory event index underestimates severity of sleep apnea compared to apnea-hypopnea index. Sleep Adv. 2023 Dec 22;5(1):zpad054. doi: 10.1093/sleepadvances/zpad054. PMID: 38264141; PMCID: PMC10805527.
18 Malhotra RK, Kirsch DB, Kristo DA, Olson EJ, Aurora RN, Carden KA, Chervin RD, Martin JL, Ramar K, Rosen CL, Rowley JA, Rosen IM; American Academy of Sleep Medicine Board of Directors. Polysomnography for obstructive sleep apnea should include arousal-based scoring: an American Academy of Sleep Medicine position statement. J Clin Sleep Med. 2018;14(7):1245–1247.
19 Pitkänen M et al. Temporal and sleep stage-dependent agreement in manual scoring of respiratory events. J Sleep Res. 2025;34(3):e14391.
20 Magalang UJ et al; SAGIC Investigators. Agreement in the scoring of respiratory events and sleep among international sleep centers. Sleep. 2013;36(4):591–596.
21 Alvarez-Estevez D, Fernández-Varela I. Large-scale validation of an automatic EEG arousal detection algorithm using different heterogeneous databases. Sleep Med. 2019;57:6–14.
22 Brink-Kjaer A et al. Clin Neurophysiol. 2020.
23 Performance based on clinical validation studies included in the clinical evaluation supporting CE marking under EU MDR (2017/745); data also submitted in FDA 510(k) K252330.
24 Parekh A, Kam K, Wickramaratne S, Tolbert TM, Varga A, Osorio R, Andersen M, de Godoy LBM, Palombini LO, Tufik S, Ayappa I, Rapoport DM. Ventilatory Burden as a Measure of Obstructive Sleep Apnea Severity Is Predictive of Cardiovascular and All-Cause Mortality. Am J Respir Crit Care Med. 2023 Dec 1;208(11):1216-1226. doi: 10.1164/rccm.202301-0109OC. PMID: 37698405; PMCID: PMC10868353.
25 Lechat B, Eckert DJ. Ventilatory Burden: Development of a New Approach to Better Quantify Obstructive Sleep Apnea Severity and Its Impacts. Am J Respir Crit Care Med. 2023 Dec 1;208(11):1153-1155. doi: 10.1164/rccm.202310-1718ED. PMID: 37878826; PMCID: PMC10868366.
26 Labarca G, Vena D, Hu WH, Esmaeili N, Gell L, Yang HC, Wang TY, Messineo L, Taranto-Montemurro L, Sofer T, Barr RG, Stone KL, White DP, Wellman A, Sands S, Redline S, Azarbarzin A. Sleep Apnea Physiological Burdens and Cardiovascular Morbidity and Mortality. Am J Respir Crit Care Med. 2023 Oct 1;208(7):802-813. doi: 10.1164/rccm.202209-1808OC. PMID: 37418748; PMCID: PMC10563185.
27 US Food and Drug Administration. Examples of Real-World Evidence Used in Medical Device Regulatory Decisions, Fiscal Years 2020-2025. April 2026. https://www.fda.gov/media/191805/download?attachment

Regulatory Information

DeepResp® is FDA-cleared medical software (K241960, K252330). DeepResp (Licence No. 113803) is a Class II Single Medical device licensed by Health Canada for professional use. DeepResp holds CE marking under MDR.

For more information on these devices, including their intended use, contraindications, and instructions for use, please consult the manufacturer's documentation at noxmedical.com/downloads.

All Artificial Intelligence analysis results must be reviewed by a certified technologist or a physician prior to diagnosis.

nox MEDICAL