Detailed Ethanol Sensor Characterization

This supplemental detailed description of the ethanol sensor characterization references Fig. 3 of the main article. The amperometric response of the developed biosensors was characterized *in-vitro* for an alcohol concentration range of 0.014 – 9.51 mM (Fig. 3a). The sensor response was monitored for successive additions of alcohol. The oxidation current increased corresponding to the change in alcohol concentration which reached a steady-state value within approximately 35 s. As shown in the Fig. 3a inset, the biosensor displayed a linear response to the alcohol concentration within the range of 0.014 to 3.67 mM. The linear equation between current (y in µA) and alcohol concentration (x in mM) was $y = (0.9206 \pm 0.012)x + (0.0569 \pm 0.0092)$ (n=3) with a correlation coefficient (R) of 0.996. From the slope of the linear range, the sensitivity was 7.32 ± 0.0989 µA·mM^{-1}·cm^{-2}. The sensor detection limit (*DL*) was estimated as 1.7 µM using the equation $DL = 3SD_B/m$, where SD_B is the standard deviation of the blank signal and m is the slope of the calibration curve.

The biosensor repeatability was evaluated by comparing the sensitivities determined from five different calibration curves over the entire linear range of the biosensor (0.014 -3.67 mM). The biosensor exhibited excellent repeatability with a mean sensitivity value of 7.25 µA·mM^{-1}·cm^{-2} and relative standard deviation (RSD) of 3.95%. In addition, the fabrication reproducibility was examined by evaluating the sensitivity within the linear range for five different sensors prepared independently on different days. From this analysis, the mean sensitivity value was 7.32 µA·mM^{-1}·cm^{-2} with a RSD of 4.84%, which indicates a moderately low sensor to sensor variation for the described fabrication process. Short term storage stability of the biosensor was also studied by dry storing the sensors at 4 °C over a 5 day period. In these studies, it was observed that the biosensors retained about 96.6% of their original response with no significant decrease in signal quality.

The panel of fouling agents was chosen for specificity testing to represent interference from possible mixing of sweat stimulant in sweat sample (1% w/w carbachol), as well as sweat born interferants in similar analytes, lipids and large proteins (1 g/L glucose, 50 mg/L stearic acid and 10 g/L BSA, respectively). Note that these concentrations represent an amplified worst-case scenario. The 1% w/w concentration of carbachol is used to fabricate the stimulant gel and would never be subjected to the sensor after dilution in sweat sample (under normal operational circumstances), and the sweat born foulant concentrations are far above those physiologically found in human sweat. Results depicting the sensor response to each foulant individually are shown in Fig 3b. The large concentrations of both carbachol and BSA cause small but noted rises in the signal when used alone but are relatively insignificant to the 1 mM ethanol response (a low clinically relevant concentration of sweat ethanol). Fig 3c depicts a demonstration where all of the fouling agents are mixed into one solution (at the same concentrations), then spiked with 1 mM ethanol. Again, a sensor response is observed over pure PBS but when spiked with the relatively low concentration of 1 mM ethanol, a significantly distinguishable signal is still achieved over the interference. While a large signal decrease can be seen in the ethanol response when combined with all the fouling agents (*e.g.*, Fig 3b to 3c), again note that the interferents are present at abnormally high concentrations.

Detailed Description of Difficulties with Iontophoretic Sweat Stimulation and The Sensor Integration

In-vivo tests were initially performed by wetting the skin-adjoining contact material C1 (Fig. 1, 2) with pure DI water and waiting for 1 hour before iontophoresis to test the membrane isolation performance (hereafter referred to as Protocol 1). This Protocol would represent more closely the conditions if repeated iontophoretic simulation were needed to provide >24 hours of sweat access, which based on our previous studies with carbachol sweat stimulant could be needed for some test subjects. Ultimately, we found Protocol 1 to be unreliable, as will be detailed further below. Therefore, instead the skin-adjoining contact material C1 was wetted with 1% carbachol in DI water and immediate stimulation used to obtain the final test subject data (hereafter referred to as Protocol 2).

Elaborating on the challenges with Protocol 1, weak sweat responses were obtained using Protocol 1 (<0.5 nL/min/gland) which is a two-fold problem caused by dosing constraints and hex wick dimension requirements. First, the applied iontophoresis current and duration were inadequate for the significant increase in resistance to iontophoretic current flow presented by the membrane isolation material C2 (Fig. 1, 2). Iontophoresis parameters (*i.e.*, applied current and duration) were limited to those for a standard Wescor Nanoduct dosing which assumes stimulant gels directly applied to the skin. This was a safety measure in the interest of the human subjects under the direction of the Institutional Review Board agreement. Thus, iontophoresis largely drove stimulant into the skin-adjoining contact material C1 with lesser than required amounts making it through and into the skin. We believe that Protocol 1 is achievable with further study and optimization of dosing conditions and by further minimizing the horizontal dimensions of the hex wick B1 and isolation adhesive B2 (Fig. 1, 2) . However, such optimization was outside of our approved dosage ranges for human-subjects testing, and further minimization of device dimensions is beyond our lab's current capability for hand-assembled devices.

Protocol 2 simulates such a condition where the adjoining contact material C1 has been properly charged with carbachol via the first phase of an optimized iontophoresis, and a significant stimulation can be achieved at the dosing parameters allowed here. Thus, results herein provide evidence into the impact of an improved/optimized iontophoretic dosing strategy for Protocol 1. Regarding minimizing hex wick B1 dimensions, the integrated patch relies on sudo-motor axon reflex (SAR) sweating, or areas under the hex wick-skin interface must receive some partially horizontal direct stimulation. Both of these stimulation mechanisms would benefit from reducing the dimensions of both the hex-wick B1 and isolation adhesive B2. Although the two finger hex wick B1 footprint was designed to distribute stimulation around the sampling area, the center of the 3 mm fingers are >1.5 mm from any directly stimulated region, and SAR sweating is known to diminish
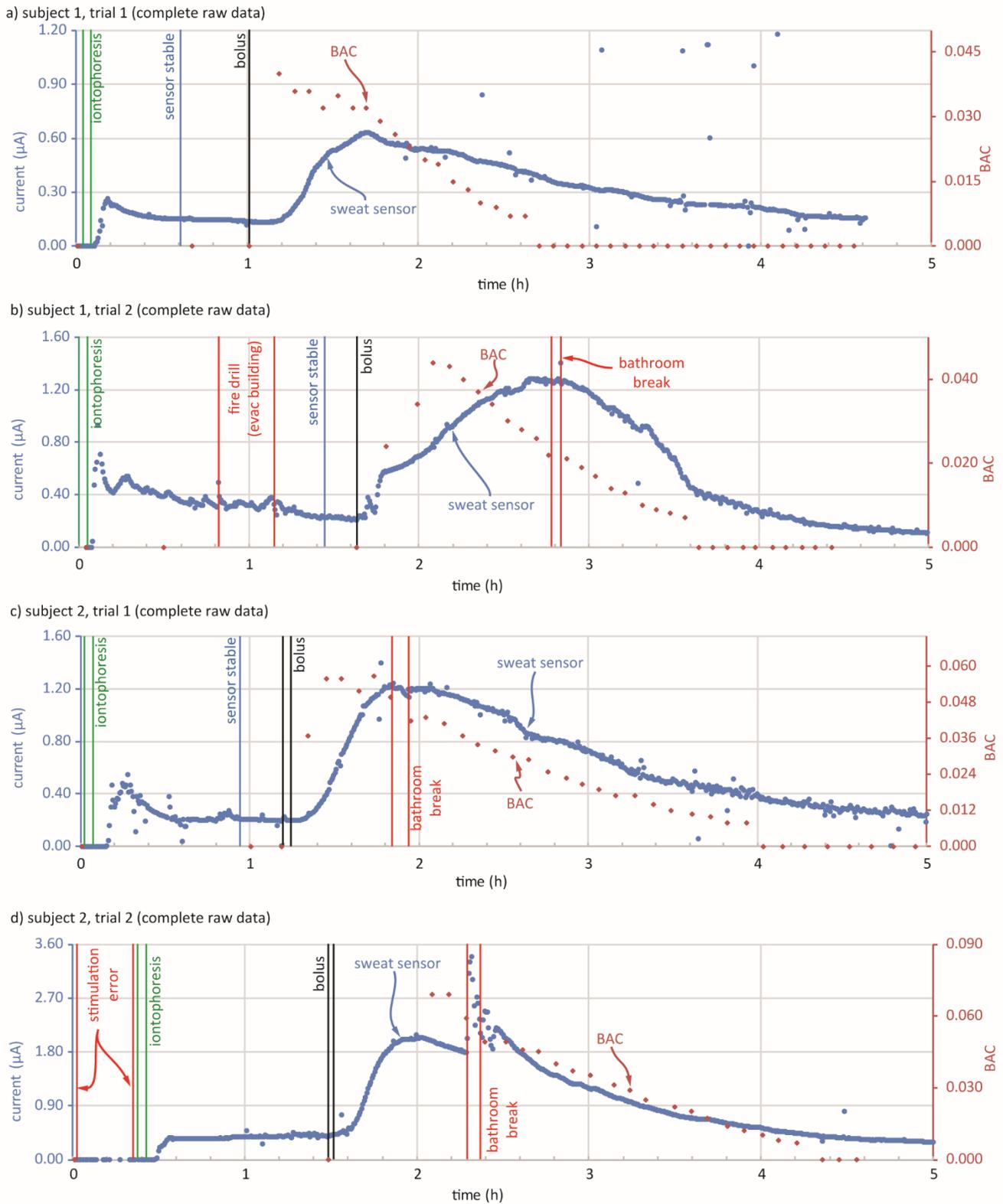
Fig. S1 Raw data from Protocol 2, used for pharmacokinetic modelling in human subjects trials (corresponding to data presented in Fig. 4 of the main manuscript) with details of significant events during testing.

significantly at that distance.[26] The easy solution to this problem is to design in more sampling fingers of thinner width (*e.g.*, four fingers of 1.5 mm width) or other optimal geometries. However, aligning such smaller and more geometrically complex components requires more accurate alignment, posing difficulties for assembly by hand in the lab, and was thus not used for these experiments. We believe that all of these stimulation issues are minor and easily resolved through dosing optimization and more sophisticated assembly methods.
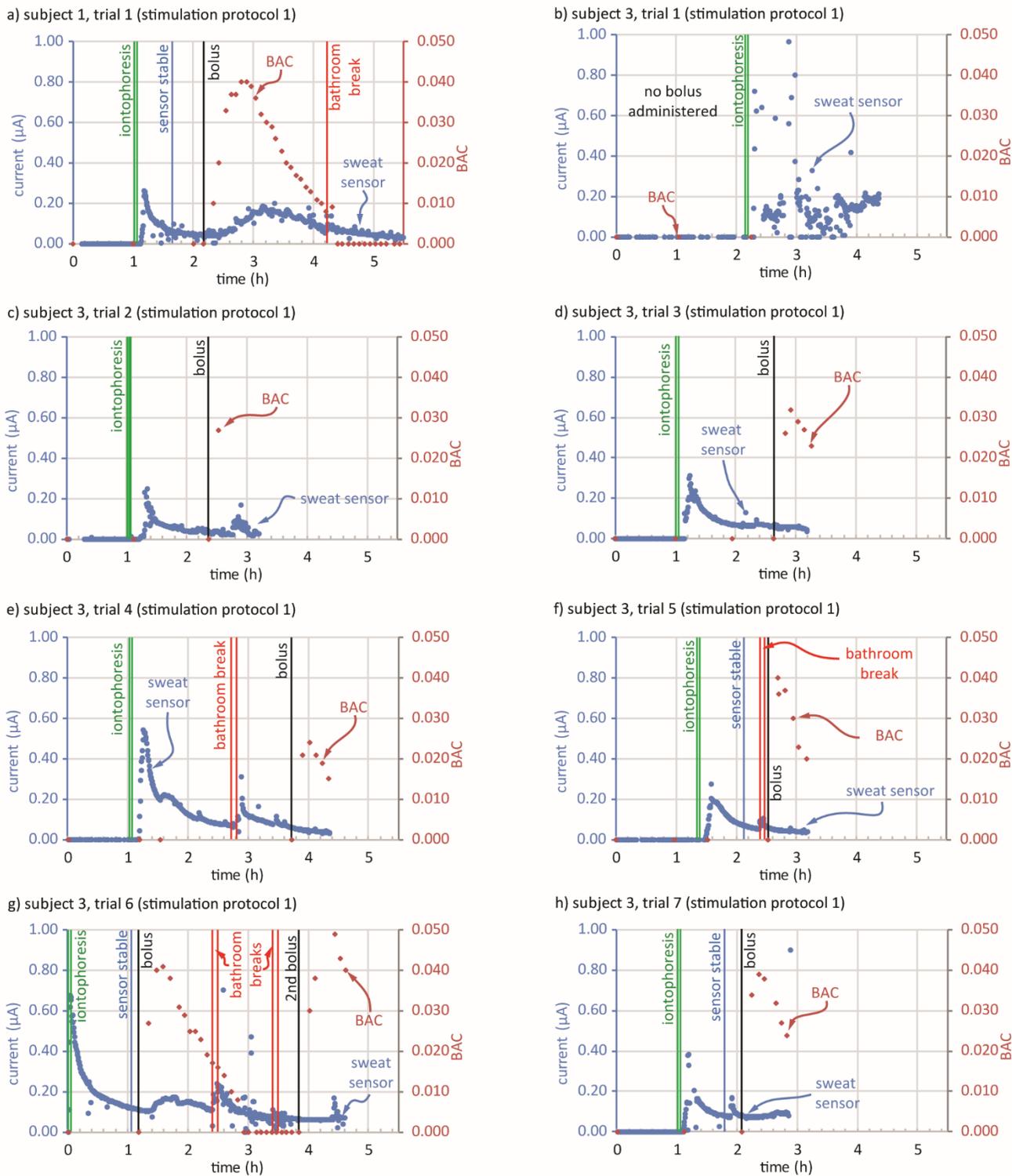
Fig. S2 Raw data of human subjects trials ran under test Protocol 1 **(a)** subject 1, trial 1, limited ethanol detection, **(b)** subject 3, trial 1, 2 h stimulation control, no bolus administered, **(c)** subject 3, trial 2, failure to detect ethanol **(d)** subject 3, trial 3, failure to detect ethanol, **(e)** subject 3, trial 4, failure to detect ethanol, **(f)** subject 3, trial 5, failure to detect ethanol, **(g)** subject 3, trial 6, limited ethanol detection, **(h)** subject 3, trial 7, failure to detect ethanol.

Even with Protocol 2, if any element in the system fails (stimulation, wicking, sensor, wick-to-sensor couple, waste pump, electrical connection, etc.) the device will fail to produce a fully successful subject test. We therefore provide further details of what tests worked or did not work for both Protocols 1 and 2, such that the reader understands the challenge level for achieving a fully successful human subject test. Some success was achieved with Protocol 1, but only two out of eight total trials successfully detected ethanol in sweat, and those successful tests were limited with noisy, low current responses (Fig. S2). However, all Protocol 1 tests showed the membrane isolation C2 method provided active control of stimulation (*i.e.*, no sweat response for 1-2 hours on skin followed by confirmed sweat response within 10 min after iontophoresis). Interestingly, significant differences in results were observed between human subjects indicating physiological variation is also likely a concern. Subject 1 was tested only once with Protocol 1 (Fig. S2a) and twice with Protocol 2 (Fig. S1a and b, Fig.
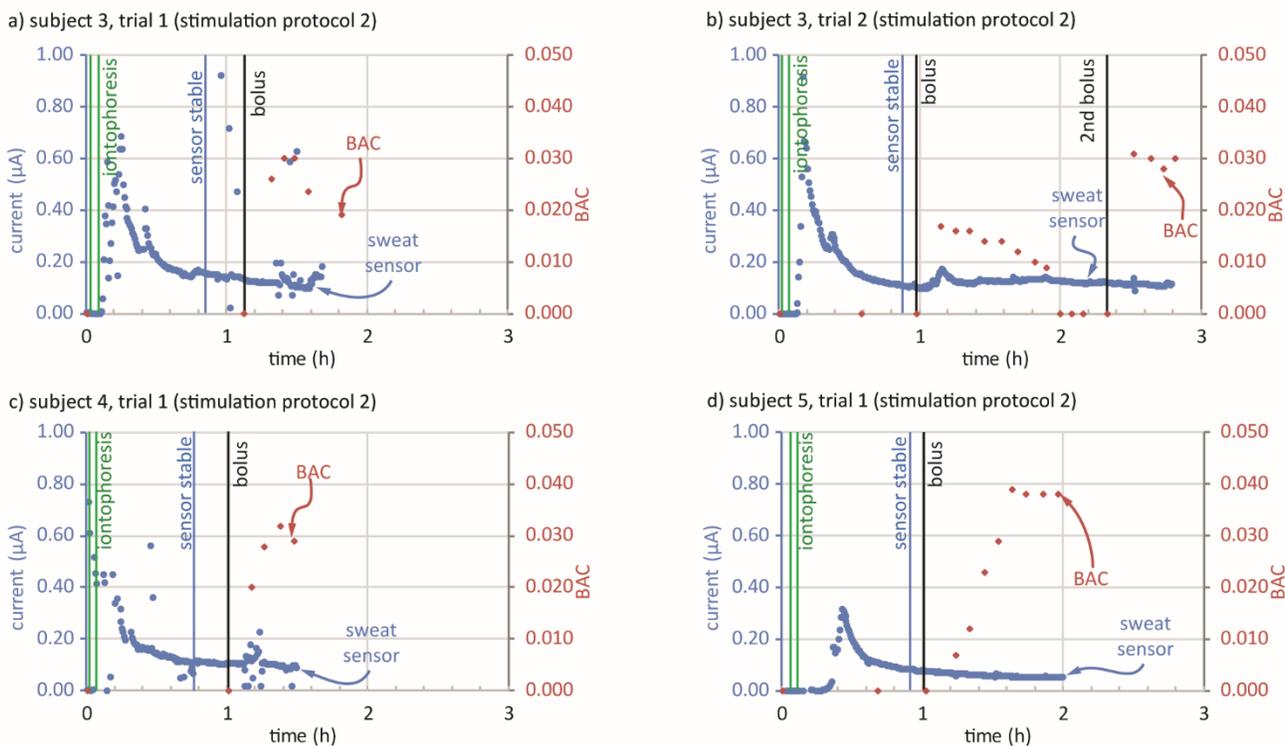
Fig. S3 Raw data of human subjects trials ran under test Protocol 2 and resulting in unsuccessful detection of ethanol in sweat **(a)** subject 3, trial 1, **(b)** subject 3, trial 2, **(c)** subject 4, trial 1, **(d)** subject 5, trial 1.

4a and b of the main article). Each one of these was a success and it is notable that Subject 1 exhibited the best results for Protocol 1. Subject 2 was only tested twice with Protocol 2 which both resulted in successful alcohol monitoring (Fig. S1c and d, Fig. 4c and d of the main article). Conversely, Subject 3 was tested seven times using Protocol 1, exhibiting an extremely low current response to alcohol in only a single trial (Fig. S2g) and failing to detect ethanol in all six other trials (Fig. S2b-f and h). Subject 3 was tested twice and Subjects 4 and 5 tested once each with Protocol 2 without success in monitoring ethanol (Fig. S3a-d). Another attempt was made on Subject 3 which provided stimulation by directly applying stimulant gels to the skin, then placing the patch directly over the stimulated area (Fig. S4). Here the sensor responded to ethanol but with a very low and noisy signal, despite confirming a significant sweat response (~5 nL/min/gland) and sample delivery to the sensor after the test. From the observations, it is speculated that Subject 1 and 2 have a stronger response to carbachol than Subjects 3-5.

We are currently unable to determine the causes of some of the above failures, and a larger human-subjects trial is suggested along with professionally (industry) fabricated devices. We speculate that many issues can be attributed to the non-ideal DropSens electrode platform. First, the large electrodes produce a large sensor footprint area and the covering dielectric used to define that area forms a 20-30 μm deep "bowl" around the sensor. This creates an unnecessarily large ~800 nL dead volume between the wick and the sensor that requires a large volume of fumed silica film (which reduces the dead volume to ~360 nL). Secondly, the thick substrate requires the hex wick to bend significantly as it is routed off of the skin and the rigidity causes it to pry the patch away from the skin when the human subject moves. These issues affect the quality of contact at the critical skin-wick and wick-sensor interfaces and also make electrical connection to the electronics difficult, contributing to noisy signals. This was most obvious when the subject needed to move considerably to use the restroom as annotated in Fig. S1 and S2. The sensor was disturbed so much in some cases that it required ~30 min to resettle. The same behavior could be observed if the sensor was disconnected and then reconnected to the electrochemical analyzer (open circuited then close circuited), reinforcing the theory that the bulk of the noise was due to the difficult electrical connection. The commercially available electrode platform was used here simply to demonstrate the robust applicability of the presented integrated system design to any reasonable sensor platform. Development of sensors on custom platforms (*e.g.*, thin, flexible substrates with optimized dimensions) would give obvious benefits that will certainly enhance performance and more reliable immobilization of the patch on the skin.
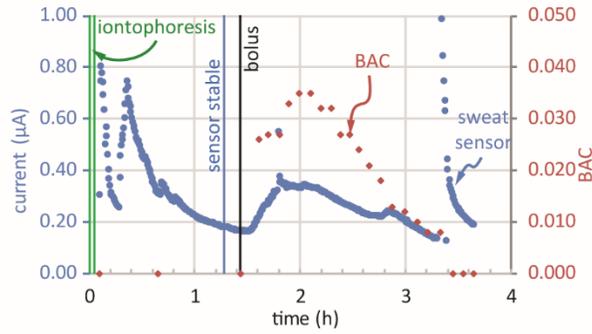
Fig. S4 Raw data from Subject 3 after directly stimulating over the collection area before placing the patch.

Details for Correlation Calculations

Correlation of the ethanol sensor data was performed with respect to BAC data points for both the model data and raw data separately. A R^2 value was calculated for each representing the amount of variation in BAC that is explained by the sensor data. Because the relation between the BAC and sensor measurements showed a large variation due to the delay in signal (see Fig. S7), signal peak correction was performed upon all data shown in Fig. 4. The modeled and raw peak-corrected data is shown below (Fig. S5 and S6). Fits are represented by R^2 values and include a 95% confidence interval. The 95% confidence interval for R^2 is calculated based on a bootstrap method where the original data was sampled 500 times with replacement and calculates the R^2 for each resampled data (ESI Table S5). The variation among these R^2 values are the basis for the confidence interval estimation. Finally, for all linear fits the simple Pearson correlation can be found by taking the square root of R^2 (ESI Table S6). Once signal peaks were corrected, the R^2 values greatly improved (Fig. S7) and showed a large linear relationship for all tests except Fig. 4a (exponential relationship, cause currently unknown). While these blood-sweat correlations show a strong relationship for the decrease in ethanol concentrations, future improvements in biosensing and predictive algorithms could be employed such that all data could be modeled (concentration rise, peak, decrease).
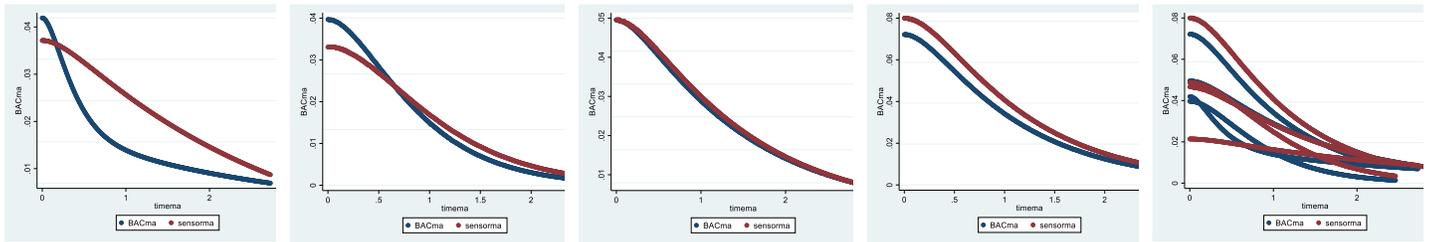
Fig. S5 Plots of peak corrected model data (blue BAC and red ethanol sensor) for the 2 trials ran on Subject 1 and 2 as well as all overlayed respectively.
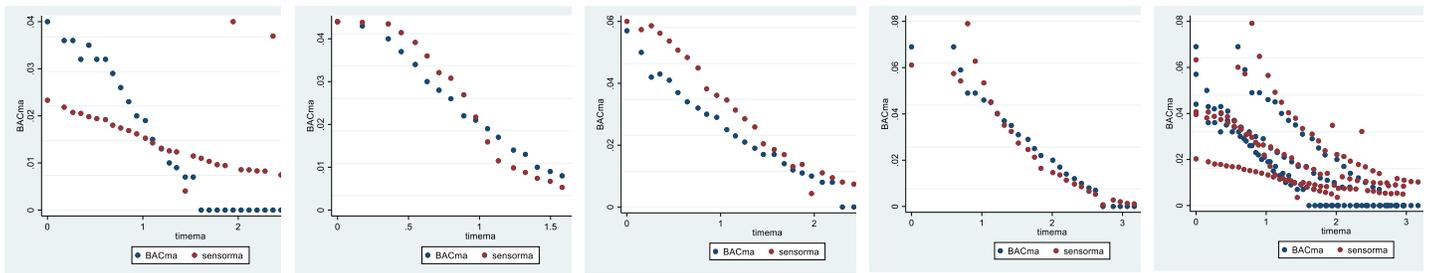
Fig. S6 Plots of peak corrected raw data (blue BAC and red ethanol sensor) for the 2 trials ran on Subject 1 and 2 as well as all overlayed respectively.

(a) uncorrected data for Model data   **(b) peak corrected data for Model data**
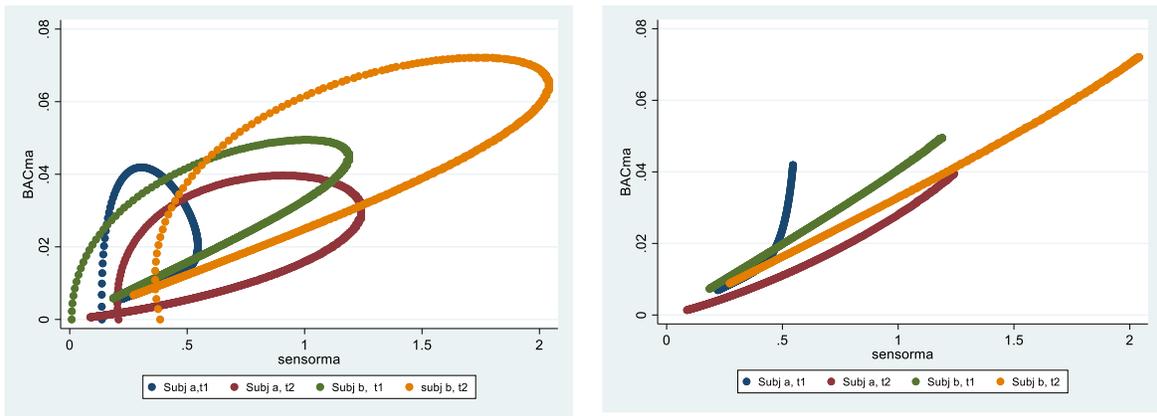
(c) uncorrected data for Raw data
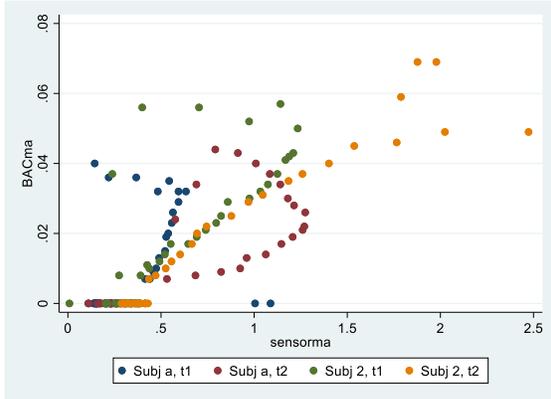
(d) peak corrected data for Raw data
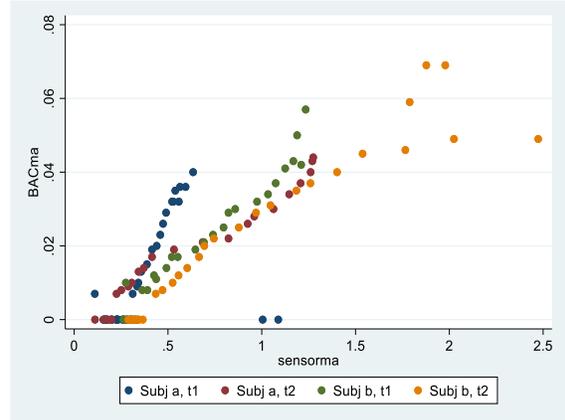
Fig. S7 Plots of relationship between BAC and ethanol sensor before and after peak correction for both model and raw data.