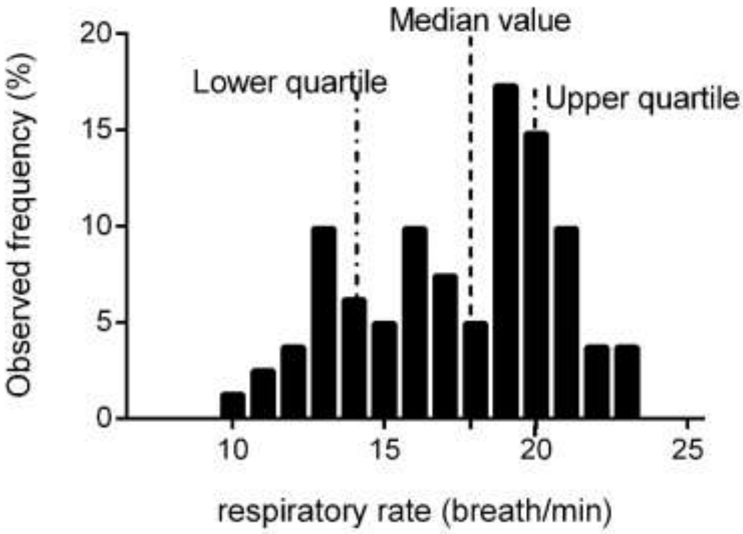
Appendix_1

Construction of an empirical population to allow prediction of a repeated observation.

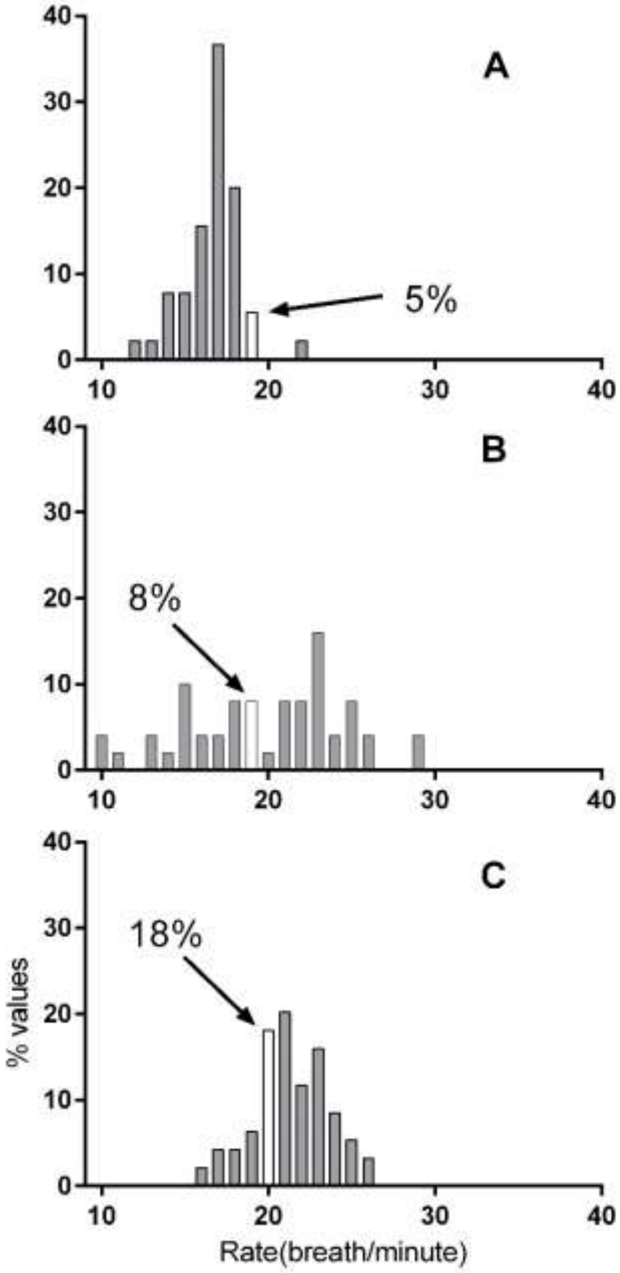
Values of respiratory rate from repeated samples from a single patient are shown below as a distribution histogram: (we have used the subject shown in figure 1B of the paper)



This distribution is based on 81 random samples. The upper quartile value is a respiratory rate of 20 breath min^{-1} . Thus if a single observation of a value of 20 breath min^{-1} were found, *in this patient,* then a repeat observation would be about 75% more likely to be less than 20, and the chance of the rate being either 19, 20, or 21 would be 42% (the sum of those % rate values shown in this figure).

In general, the distributions of rates observed from the patients we sampled show that the likelihood of obtaining an exact particular value again, if a measurement were repeated, is small. The likelihood of a repeat 30 second measurement *in a particular patient* having the *same* value as a previous measurement is less than 30%. However, we can only make these inferences when we already know the individual distributions of observed rates, *in each particular patient.*

Many breathing rate values we measured can be found in the distributions of values from different patients. For example, three patients in adjacent beds might all, at some time, have had a rate of 19 breath min^{-1}, although their overall respiratory rate distributions differ substantially. Thus, presented with a “new patient” who had a single respiratory rate measurement of 19, this single value might be observed, at some time, in any of these 3 patients.



Rate estimates for three patients, A, B, and C. In each patient, the proportion of observations of the rate value of 19 are indicated.

The rate value 19 might be observed at some time in any of these three patients: the likelihood of such an observation in patient A is small, more in B, and substantial in C.

Another measurement, if it were from patient A, would be more likely to be less than 19. A repeat measure from patient B could be either greater or less than 19, whereas if the repeat were from patient C, a value greater than 19 would be more likely.

We can estimate the likelihood of repeat observations, by using all our sample observations to make an empirical likelihood distribution. The principle is explained below with a simple example that uses only 4 patients, a limited range of possible rates, and an artificial, symmetrical, and limited distribution of rate observations. Our actual model contained 25 patients and 51 observed values of respiratory rates, and the distributions within subjects varied substantially (see figure 2 of the paper)

	Patients				
Rate (breath/min)	A	B	C	D	Sum of probabilities
14	2	0	0	0	2
15	3	2	0	0	5
16	8	3	2	0	13
17	18	8	3	2	31
18	38	18	8	3	67
19	18	38	18	8	82
20	8	18	38	18	82
21	3	8	18	38	67
22	2	3	8	18	31
23	0	2	3	8	13
24	0	0	2	3	5
25	0	0	0	2	2
	100	100	100	100	

The distribution of rate observations found each of four example patients, A to D, expressed as %. These are the *within patient* probabilities of finding a specific rate. In this example, these distributions are set to have the same shape, but have different central values, of 18, 19, 20 and 21 breath min^{-1}. The sum of the probabilities of observing each rate, for the four patients, is given in the column on the right: thus the rates of 19 and 20 have a sum of 82: each of these rates is likely to be found in 20·5% of observations taken from this group of 4 patients (i.e. 82/4 %)

Rate	Within Patients A	B	C	D	Between Patients A	B	C	D	Sum of probabilities
14	2	0	0	0	100	0	0	0	100
15	3	2	0	0	60	40	0	0	100
16	8	3	2	0	62	23	15	0	100
17	18	8	3	2	58	26	10	6	100
18	38	18	8	3	57	27	12	4	100
19	18	38	18	8	22	46	22	10	100
20	8	18	38	18	10	22	46	22	100
21	3	8	18	38	4	12	27	57	100
22	2	3	8	18	6	10	26	58	100
23	0	2	3	8	0	15	23	62	100
24	0	0	2	3	0	0	40	60	100
25	0	0	0	2	0	0	0	100	100

The columns added on the right express how the % probability of finding any specific rate is distributed *between* these 4 example patients. For example, in the top shaded row of the table, a rate of 14 is only observed, twice, in a single patient (patient A). If we are given a sample from one of the four patients shown here, and the value was 14, it would be 100% probable that the sample was from patient A.

In contrast, a rate of 20 is found in all the patients. (Lower shaded row) If we obtained a sample from one of these four patients, and the value was 20, the probability that the sample had come from patient A is only 10%. In contrast, the probability that the sample would have come from patient C is 46%.

Given a specific observed value from this population, we combine the likelihood that this observed value has come from a specific patient (which expresses the *between* patient variation for specific rates) and the likelihood of the values observed within each patient (the *within* patient variation) to predict the likelihood of the value of any repeat observation.

For example, if we observe a rate of 14, which would only be found if the sample had been taken from patient A, a repeat observation of the same individual would be drawn only from this patient's spectrum of rate values (dotted box, column A). The most likely rate found in a repeat sample would be 18.

On the other hand, if we observe a rate of 20, then we have to consider all the patients A to D to be potential sources for the observation. Patient C is the most likely patient to have given a rate of 20.

We can now extend this table to show predictions of a repeat observation of 20 breath/minute drawn from the matrix: see table 3.

Table 3

Rate	Within Patients A	B	C	D	Between Patients A	B	C	D	Sum of probabilities	Patient contribution matrix (based on an observed value of 20) A	B	C	D	Sum of probabilities
14	2	0	0	0	100	0	0	0	100	0·2	0·0	0·0	0·0	0·20
15	3	2	0	0	60	40	0	0	100	0·3	0·4	0·0	0·0	0·73
16	8	3	2	0	62	23	15	0	100	0·8	0·7	0·9	0·0	2·37
17	18	8	3	2	58	26	10	6	100	1·8	1·8	1·4	0·4	5·34
18	38	18	8	3	57	27	12	4	100	3·7	4·0	3·7	0·7	12·02
19	18	38	18	8	22	46	22	10	100	1·8	8·3	8·3	1·8	20·20
20	8	20	38	18	10	22	46	22	100	0·8	4·0	17·6	4·0	26·29
21	3	8	18	38	4	12	27	57	100	0·3	1·8	8·3	8·3	18·73
22	2	3	8	18	6	10	26	58	100	0·2	0·7	3·7	4·0	8·51
23	0	2	3	8	0	15	23	62	100	0·0	0·4	1·4	1·8	3·59
24	0	0	2	3	0	0	40	60	100	0·0	0·0	0·9	0·7	1·59
25	0	0	0	2	0	0	0	100	100	0.0	0.0	0.0	0.4	0·44
	100	100	100	100										

The "contribution matrix" shows likelihoods of finding any rate, if a second observation is taken from any patient who could have provided a specified rate value.
Consider an observation of the rate 20. (Horizontal dashed box in the "between patients" columns in the centre). A value of 20 could have come from any of the four patients, A, B, C, or D. The likelihood that it comes from patient A is 10%, shown in the heavy dotted box of the central row.
If a second observation were made of patient A, the likelihood of the possible values that would be drawn are shown in the distribution on the left (vertical dashed box).

Consider the probability of obtaining the sample value of 21. The probability that we would obtain a random sample of 21 from patient A is 3%: see the circle in the dashed vertical box. If the repeat observation we make has a 10% likelihood of being from patient A, then the possibility of an observation of the rate 21, from patient A, becomes 10% of 3%, i.e. 0.3%: this is entered in the black outlined box of the Patient contribution matrix (four columns on the right). The four columns on the right contain the probabilities of each patient, A, B, C, and D yielding a repeat observation rate, when the original rate was 20.

The last column (shaded) sums the probabilities of the different rates being found in all 4 patients, if the rate of 20 has been observed in a first sample, and the possibility of obtaining this rate has been drawn from values available from these four patients.

In the paper, figure 3 shows values that might be found if rates of 12, 20, and 25 were made, and then these observations were repeated. The open column indicates the chance that the

first observation would be replicated in a further random sample, and the grey columns the chances of observing values within 2 breath min^{-1} of the initial observation. The likelihood of close repeat measures (i.e consistency) is greater if the observation duration is greater.
For 30 second samples, and an initial rate of 12, then 40% of repeated measures would be 14 breath min^{-1} or more, which is well within the “normal range”. Similarly, for an initial rate of 25, 40% of repeat values would be less than 23 breath min^{-1}. Overall, less than 50% of repeat 30 second samples lie within 2 breath min^{-1} of the original value. By extending the sample time to 2 minutes, overall about 60% of repeat measurements will then fall within 2 breath min^{-1} of the original.