An improved statistical methodology to estimate and analyze impedances and transfer functions

DOUGLAS CURRAN-EVERETT, YIMING ZHANG, M. DOUGLAS JONES, J. R., AND RICHARD H. JONES

Departments of Pediatrics and of Preventive Medicine and Biometrics, School of Medicine, University of Colorado Health Sciences Center, Denver, Colorado 80262

Curran-Everett, Douglas, Yiming Zhang, M. Douglas Jones, J. R., and Richard H. Jones. An improved statistical methodology to estimate and analyze impedances and transfer functions. J. Appl. Physiol. 83: 2146–2157, 1997.—Estimating the mathematical relationship between pulsatile time series (e.g., pressure and flow) is an effective technique for studying dynamic systems. The frequency-domain relationship between time series, often calculated as an impedance (pressure/flow), is known more generally as a frequency-response or transfer function (output/input). Current statistical methods for transfer function analysis 1) assume erroneously that repeated observations on a subject are independent, 2) have limited statistical value and power, or 3) are restricted to use in single subjects rather than in an entire sample. This paper develops a regression model for transfer function analysis that corrects each of these deficiencies. Spectral densities of the input and output time series and the cross-spectral density between them are first estimated from discrete Fourier transforms and then used to obtain regression estimates of the transfer function. Statistical comparisons of the transfer function estimates use a test statistic that is distributed as χ². Confidence intervals for amplitude and phase can also be calculated. By correctly modeling repeated observations on each subject, this improved statistical approach to transfer function estimation and analysis permits the simultaneous analysis of data from all subjects in a sample, improves the power of the transfer function model, and has broad relevance to the study of dynamic physiological systems.

discrete Fourier transform; frequency-domain regression; frequency-response function; mixed-effects model; spectral analysis

ESTIMATION OF THE MATHEMATICAL RELATIONSHIP between pulsatile time series is a time-honored approach to the study of dynamic systems. Using this strategy, physiologists have explored respiratory mechanics (11, 15, 23, 30, 39), pulmonary ventilation (29, 33), cardiovascular function (3, 7, 24, 27, 32), and cardiorespiratory regulation (4, 25, 34, 35). The relationship between time series is calculated most often in the frequency domain (5, 24); this entails first transforming the time series, originally functions of time, into their equivalent Fourier coefficients, written as functions of frequency. When two time series represent the input and output of a dynamic system, the mathematical relationship between them is often designated the transfer function (h), defined in terms of frequency (f) as h(f) = output(f)/input(f), where the output and input functions are the Fourier transforms of the original time series. When input is flow and output is pressure, the transfer function is called impedance. A transfer function is a complex expression, with real and imaginary parts, that can be described by amplitudes (output amplitude divided by input amplitude) and phase angles (timing of output with respect to input).

In experimental situations, transfer function analysis is complicated by between-subject variability as well as by random measurement error. Current approaches to transfer function analysis 1) assume erroneously that repeated observations on a subject are independent, 2) have limited statistical value and power, or 3) because they are unable to account for between-subject variability, are restricted to use in single subjects rather than in an entire sample. This paper develops a regression model for transfer function analysis that corrects each of these deficiencies and has broad relevance to the analysis of dynamic physiological systems. Before we derive and illustrate this improved methodology, we review previous approaches to applied transfer function analysis.

APPLIED TRANSFER FUNCTION ANALYSIS

Description. Many studies (3, 7, 11, 15, 23, 26, 28, 34, 42) simply describe the effect of an experimental intervention on a transfer function; they report, for example, changes in the frequencies at which maxima in impedance amplitude occur. But to draw conclusions about an underlying population, one must use more than rudimentary description: one must employ the inferential techniques of confidence intervals and significance tests (38).

Usual indexes of between-subject variability. In most physiological studies (see Ref. 24), between-subject variability is estimated, first by deriving the amplitude and phase angle of a transfer function for all subjects and then by calculating SDs. This standard approach of handling between-subject variability fails to account for the fact that repeated observations on a subject (e.g., observations made during baseline and then during an experimental intervention) are correlated (21). Because of this correlation, the true error variabilit-
variabilities (21). The Appendix shows how correlation 
SDs of amplitude and phase underestimate the true 
ity is underestimated, and the reported values for the 
Separate analyses of amplitude and phase. Some 
researchers (4, 35, 40) analyze the amplitude and phase of a 
transfer function as if they were unrelated vari-
bles. But the real and imaginary parts of a transfer function, from which amplitude and phase are derived, 
are determined simultaneously. Therefore, the com-
ponents of a transfer function, either its real and imagi-
nary parts or its amplitude and phase, must also be 
analyzed simultaneously. There is a quantitative ben-
efit: the simultaneous analysis of jointly derived data improves statistical power (10).

Derivation and analysis of analog parameters. Be-
cause transfer function estimation simply reduces a 
relationship between pulsatile time series to its mathe-
matical form, the correspondence of a transfer func-
tion to physical characteristics of a system is unclear. 
To circumvent this limitation, some investigators (15, 
39, 42) first obtain an estimate of the transfer function and 
then, from that estimate, derive analog parameters that represent general characteristics (e.g., total compli-
ance) or specific components (e.g., lung tissue elas-
tance) of the system. It is the estimates of the analog 
parameters, rather than the estimate of the transfer 
function, that are subjected to statistical analysis. 
Although this approach gives physiological meaning to 
a transfer function, it does have a statistical drawback: 
the estimates of the analog parameters are correlated 
(38). This means that the analysis of only one (prese-
lected) parameter is valid: the statistical outcomes of 
the remaining analyses will be related to the outcome of 
the initial analysis.

Time-domain analysis. The time-domain technique 
of bivariate autoregression has been used also to esti-
mate a transfer function (22). Bivariate autoregres-
sions (20) can be fit to the original time series by using 
the Yule-Walker equations, which define the relation-
ships between the covariance functions, the cross-
covariance function, and the autoregression matrices 
(18, 41). The resulting bivariate autoregressive esti-
mates are related to the autocorrelation and cross-
correlation functions, often used in physiology to esti-
mate a transfer function (24, 25, 32). From the estimated 
autoregressive matrices, estimates of the spectral densi-
ties of an input and output and the cross-spectral density between them can be calculated; it is from these 
autoregressive spectra that a transfer function can be 
computed (see Preliminary Data Analysis). The ad-
vantage of autoregression is that it can produce smooth 
spectral estimates of a transfer function. The disadvan-
tage is that the autoregressive estimates are correlated 
across frequency (20, 37); therefore, the statistical properties required to compare transfer functions at 
specific frequencies are unknown.

Regression techniques. Since the 1940s, regression 
analysis in the frequency domain has been used in 
transfer function estimation (see Ref. 6). In this ap-
proach, the real and imaginary parts of the Fourier 
transforms of two time series are used to obtain a 
regression estimate of the transfer function (6, 16, 31, 
37). The classic regression model (see Eq. 3) assumes, 
however, that estimates of the transfer function during 
a given condition are virtually identical for all subjects. 
This is unrealistic; physiological transfer function esti-
mates vary considerably among subjects (7, 8, 28, 32, 
34, 42). Indeed, it was explicitly because of between-
subject differences that O’Rourke and Taylor (Ref. 28, 
p. 368–369) presented only individual impedance spec-
tra: “Although statistical analysis of a large number of 
experimental results is desirable, it is obvious that 
features of the impedance curves can be lost if data are 
pooled from [many subjects] . . . .”

Because it does not allow for between-subject variabil-
ity, the classic regression model can estimate a transfer 
function in single subjects only (37); this confines 
statistical and biological inferences to single individu-
als rather than to an underlying population. The 
methodology we develop here extends this classic regres-
sion approach in such a way that all subjects in a 
sample can be analyzed simultaneously. This statistical 
advantage has important theoretical and applied 
advantages. In the next section, we review the experi-
mmental problem that motivated the development of this 
methodology.

ORIGINAL EXPERIMENTAL PROBLEM

The development of this methodology was driven by 
our interest in the effects of systemic circulatory pertur-
bations on dynamic properties of the cerebral circula-
tion (see Ref. 8). In this context, we considered systemic 
arterial blood pressure to be the input and cerebral 
cortical blood flow to be the output of a linear vascular 
system. In 10 rabbits, arterial blood pressure, mea-
ured by high-fidelity transducer catheter, and cortical 
flow, measured by laser-Doppler flowmeter, were 
sampled for 1 min during two conditions: baseline 
and combined hypoxia-hypercapnia. Our experimental ob-
gective was to estimate the transfer function between 
arterial blood pressure and cortical blood flow during 
each condition by using data from the entire sample. To 
do this, we selected a 20-s segment from each time 
series, giving 800 pairs of observations (sampling inter-
val 0.025 s) for each subject during each condition (Fig. 
1). The individual transfer functions derived from each 
of these bivariate time series could be estimated reli-
ably at the frequencies of heart rate and its first har-
monic.

For the frequency of heart rate, Fig. 2 
depicts the between-subject variability inherent to 
transfer function estimation. The methodology that 
follows was devised to incorporate between-subject variability into the regression model for the transfer

---

1 Although our statistical methodology can be used to estimate a transfer function at any frequency of interest, we use only the frequency of heart rate to illustrate the procedure.
function between arterial blood pressure and cortical blood flow.

**PRELIMINARY DATA ANALYSIS**

Before we derived the regression model for the transfer function, however, we verified an important statistical assumption: that the error term of the model, i.e., the spectral density of the residual series, was nearly constant around the frequencies of heart rate and its first harmonic. We did this using bivariate autoregression (20), identifying the order of the model by Akaike's information criterion (1, 2). The estimated autoregressive spectra of pressure and flow (\(s_{xx}^\hat{}\) and \(s_{yy}^\hat{}\), respectively) showed striking peaks at the frequencies of heart rate and its harmonics (Fig. 3). The estimated spectral density of the residual series was then derived in a two-step procedure analogous to that used in regression analysis. First, the transfer function estimate \(h^\hat{}\) was calculated from the bivariate autoregressive spectral estimates as

\[
h^\hat{}(f) = \frac{s_{yx}^\hat{}(f)}{s_{xx}^\hat{}(f)}
\]

where \(s_{yx}^\hat{}\) is the estimated cross-spectral density between pressure and flow.
between the flow and pressure series. Then, the estimated spectral density of the residual series ($\hat{s}_n$) was computed as

$$\hat{s}_n(f) = \hat{s}_y(f) - \hat{s}_x(f) \hat{h}(f)$$

where $\hat{s}_y$ is the complex conjugate of $\hat{s}_x$. The absence of strong peaks in the residual spectrum around the frequencies of heart rate and its first harmonic (Fig. 4) verified the requisite assumption of the regression model and enabled smoothing across bands centered at these frequencies (19). Assuming that the quantity being estimated is nearly constant within the frequency band, smoothing increases the precision of the estimate: the wider the band, the greater the precision. This strategy is valid because discrete Fourier transforms at different frequencies are virtually independent.

Next, we review time series regression in the frequency domain. As we develop the full regression model for the transfer function, we illustrate the process for the transfer function (evaluated at heart rate) between arterial blood pressure and cortical blood flow using data from Ref. 8.

### TIME SERIES REGRESSION

Hannan (16) pioneered time series regression in the frequency domain (see also Refs. 6, 31, and 37). If $x_j$ and $y_j$ represent the input and output time series, respectively, each with $n$ observations at time intervals of $\delta$, then the discrete Fourier transforms ($z_v$) of these series are

$$z_v^{(x)} = \sum_{j=0}^{n-1} x_j e^{-2\pi i j / n} \quad \text{and} \quad z_v^{(y)} = \sum_{j=0}^{n-1} y_j e^{-2\pi i j / n} \quad (1)$$

where the index $v$ is associated with frequency $f$ (Hz) by

$$f = v/(n\delta) \quad \text{and} \quad i = \sqrt{-1}.$$ Each transform $z_v$ can be written with real and imaginary parts as $z_v = \alpha_v + i\beta_v$, where

$$\alpha_v^{(x)} = \sum_{j=0}^{n-1} x_j \cos \left(2\pi v j / n\right)$$

and

$$\beta_v^{(x)} = -\sum_{j=0}^{n-1} x_j \sin \left(2\pi v j / n\right). \quad (2)$$

$\alpha_v^{(y)}$ and $\beta_v^{(y)}$ have similar equations.

The example. Because we study dynamic vascular properties at discrete frequencies, the first task is to identify the frequency of heart rate; we do this by using the spectral density of arterial pressure. Then, for each of 15 0.05-Hz frequency bands centered at heart rate, we calculate the real and imaginary components (Eq. 2) of the discrete Fourier transforms of both the input and output time series (Table 1).

Transfer function estimation. Using the real and imaginary parts (Eq. 2) of the Fourier transforms, the transfer function $h$ can be approximated in the frequency domain as

$$z_v^{(y)} = z_v^{(x)} h + z_v^{(n)} \quad (3)$$

where $z_v^{(n)}$ is the unobservable discrete Fourier transform of an additive noise series. An estimate of the transfer function can be obtained by treating Eq. 3 as a
complex regression equation. First, the spectral densities of the input and output series, $\hat{s}_x(f)$ and $\hat{s}_y(f)$, and the cross-spectral density $\hat{s}_{xy}$ between them are estimated within the frequency band centered at a relevant Fourier frequency $f$ (in the worked example, the frequency of heart rate).

$$\hat{s}_x(f) = \sum_{k=-w}^{k+w} [z_v^{(x)}]^* z_v^{(x)} = \sum_{k=-w}^{k+w} [\alpha_v^{(x)}]^2 + [\beta_v^{(x)}]^2$$

$$\hat{s}_y(f) = \sum_{k=-w}^{k+w} [z_v^{(y)}]^* z_v^{(y)} = \sum_{k=-w}^{k+w} [\alpha_v^{(y)}]^2 + [\beta_v^{(y)}]^2$$

$$\hat{s}_{xy}(f) = \sum_{k=-w}^{k+w} [z_v^{(x)}]^* z_v^{(y)} = \sum_{k=-w}^{k+w} [\alpha_v^{(x)} - i\beta_v^{(x)}][\alpha_v^{(y)} + i\beta_v^{(y)}]$$

where * denotes the complex conjugate. These estimates are smoothed periodogram estimates with uniform weights, a smoothing span of $2w + 1$, and $4w + 2$ degrees of freedom (19). Next, the real and imaginary parts of the cross-spectral density, the cospectrum $\hat{c}_{xy}$ and the quadrature spectrum $\hat{q}_{xy}$, are estimated as

$$\hat{c}_{xy}(f) = \sum_{k=-w}^{k+w} [\alpha_v^{(x)} \alpha_v^{(y)} + \beta_v^{(x)} \beta_v^{(y)}]$$

and

$$\hat{q}_{xy}(f) = \sum_{k=-w}^{k+w} [\alpha_v^{(x)} \beta_v^{(y)} - \beta_v^{(x)} \alpha_v^{(y)}]$$

Last, if the residual spectral density is nearly constant within the frequency band, an assumption verified by the preliminary analysis (see Fig. 4), then the complex regression estimate $\hat{h}$ of the transfer function $h$ can be calculated as

$$\hat{h}(f) = \frac{\hat{s}_{xy}(f)}{\hat{s}_x(f)} = \frac{\hat{c}_{xy}(f) + i\hat{q}_{xy}(f)}{\hat{s}_{xx}(f)}$$

The real and imaginary parts of Eq. 4

$$\hat{c}_{h}(f) = \frac{\hat{c}_{xy}(f)}{\hat{s}_{xx}(f)} \quad \text{and} \quad \hat{q}_{h}(f) = \frac{\hat{q}_{xy}(f)}{\hat{s}_{xx}(f)}$$

can be used to represent the estimate $\hat{h}$ with an amplitude $\hat{A}$ and a phase angle $\hat{\theta}$

$$\hat{A}(f) = \sqrt{\hat{c}_{h}^2 + \hat{q}_{h}^2} \quad \text{and} \quad \hat{\theta}(f) = \tan^{-1}(\frac{\hat{q}_{h}/\hat{c}_{h})}{\hat{c}_{h}}$$

The computational sign of $\hat{c}$ can be reversed if the sign of the exponent in the Fourier transform (Eq. 1) is reversed; in some software, the form of the transform is unclear. In transfer function applications, the timing of the output time series with respect to the input time series may be predictable: the output will lag behind the input. In impedance applications, however, the sign of the phase angle typically reverses as frequency increases. Regardless of specific application, the safest strategy is to confirm the correct computational sign of the estimated phase angle $\hat{\theta}$ by using known input and output time series.

The example. From the real and imaginary components of the discrete Fourier transforms within the 15-interval frequency band (see Table 1), we next obtained a single smoothed estimate of the transfer function (Table 2; see Transfer function estimation above). The phase estimate is meaningful only if there is high coherence between the input and output: that is, if changes in the output correspond closely to changes in the input. Based on the estimated autoregressive spectrum of squared coherence (see Fig. 3), we expect to obtain a reliable estimate of the transfer function at the frequency of heart rate. Although smoothing increases the precision of the transfer function estimate, it can also decrease squared coherence. Therefore, after we get the single transfer function estimate, we verify that squared coherence across the 15-interval frequency band remains high

$$\text{Squared coherence}(f) = \frac{\hat{c}_{xy}^2(f) + \hat{q}_{xy}^2(f)}{\hat{s}_{xx}(f)\hat{s}_{yy}(f)}$$

where $\hat{c}_{xy}$, $\hat{q}_{xy}$, $\hat{s}_{xy}$, and $\hat{s}_{yy}$ are pooled over all experimental conditions.

**RANDOM SUBJECT EFFECT**

In the preceding section, we reviewed the classic regression model (Eq. 3) that assumes, for some conditions, that estimates of a transfer function are virtually identical for all subjects. In this section, we first add a random parameter $g$ to the transfer function model, thereby transforming the classic fixed-effects model (Eq. 3) into a mixed-effects model that can handle inherent between-subject variability in the transfer

---

3 These spectral density estimates are proportional to their true spectral densities; to be a proper estimate, each would have to be multiplied by a constant so that the area under its spectral density curve estimated the variance of its respective time series.
Table 2. Transfer function estimates for 10 subjects

<table>
<thead>
<tr>
<th>k</th>
<th>( \alpha_n )</th>
<th>( \beta_n )</th>
<th>( \hat{A} )</th>
<th>( \hat{\delta} )</th>
<th>( \alpha_n )</th>
<th>( \beta_n )</th>
<th>( \hat{A} )</th>
<th>( \hat{\delta} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.0256</td>
<td>-0.1076</td>
<td>0.1016</td>
<td>-1.804</td>
<td>-0.0281</td>
<td>-0.1067</td>
<td>0.1104</td>
<td>-1.828</td>
</tr>
<tr>
<td>2</td>
<td>0.0554</td>
<td>-0.1357</td>
<td>0.1466</td>
<td>-1.183</td>
<td>0.0832</td>
<td>-0.1706</td>
<td>0.1898</td>
<td>-1.117</td>
</tr>
<tr>
<td>3</td>
<td>0.0053</td>
<td>-0.1041</td>
<td>0.1042</td>
<td>-1.520</td>
<td>0.0040</td>
<td>-0.1200</td>
<td>0.1201</td>
<td>-1.537</td>
</tr>
<tr>
<td>4</td>
<td>0.0060</td>
<td>-0.1457</td>
<td>0.1458</td>
<td>-1.530</td>
<td>0.0016</td>
<td>-0.1394</td>
<td>0.1394</td>
<td>-1.560</td>
</tr>
<tr>
<td>5</td>
<td>0.0153</td>
<td>-0.1400</td>
<td>0.1408</td>
<td>-1.462</td>
<td>0.0280</td>
<td>-0.1646</td>
<td>0.1670</td>
<td>-1.403</td>
</tr>
<tr>
<td>6</td>
<td>-0.0022</td>
<td>-0.1218</td>
<td>0.1218</td>
<td>-1.589</td>
<td>0.0086</td>
<td>-0.1213</td>
<td>0.1216</td>
<td>-1.500</td>
</tr>
<tr>
<td>7</td>
<td>0.0034</td>
<td>-0.1750</td>
<td>0.1750</td>
<td>-1.551</td>
<td>-0.0151</td>
<td>-0.1902</td>
<td>0.1908</td>
<td>-1.650</td>
</tr>
<tr>
<td>8</td>
<td>0.0476</td>
<td>-0.1015</td>
<td>0.1211</td>
<td>-1.132</td>
<td>0.0528</td>
<td>-0.1313</td>
<td>0.1415</td>
<td>-1.188</td>
</tr>
<tr>
<td>9</td>
<td>-0.0139</td>
<td>0.1565</td>
<td>0.1572</td>
<td>-1.660</td>
<td>-0.0154</td>
<td>0.1569</td>
<td>0.1576</td>
<td>-1.669</td>
</tr>
<tr>
<td>10</td>
<td>-0.0162</td>
<td>-0.2094</td>
<td>0.2100</td>
<td>-1.648</td>
<td>0.0005</td>
<td>-0.2111</td>
<td>0.2111</td>
<td>-1.569</td>
</tr>
</tbody>
</table>

Values are in ml·min\(^{-1}\)·100 g\(^{-1}\)-mmHg\(^{-1}\) for the real and imaginary parts of transfer function estimates \( \alpha_n \) and \( \beta_n \), respectively; in ml·min\(^{-1}\)·100 g\(^{-1}\)-mmHg\(^{-1}\) for amplitude \( \hat{A} \); and in radians for phase \( \hat{\delta} \). For each subject, transfer function was estimated over a band of 15 frequencies; this band was centered at subject-specific heart rate. \( k \), Subject no. Treatment is combined hypoxia-hypercapnia.
For a sample of $N$ subjects, the estimate $\hat{h}$ of the transfer functions is

$$\hat{h} = \left[ \sum_{k=1}^{N} z_k^{(x)\ast} V_k^{-1} z_k^{(y)} \right]^{-1} \left[ \sum_{k=1}^{N} z_k^{(x)\ast} V_k^{-1} z_k^{(y)} \right]$$

(7)

Because

$$V_k^{-1} = I - \frac{c^2 z_k^{(2)\ast} z_k^{(2)}}{1 + c^2 z_k^{(2)\ast} z_k^{(2)}}$$

in the single transfer function case (i.e., when the vector $h$ is a scalar: $z_k^{(2)} = z_k^{(y)}$), the estimate reduces to

$$\hat{h} = \left[ \sum_{k=1}^{N} \frac{z_k^{(x)\ast} z_k^{(x)}}{1 + c^2 z_k^{(x)\ast} z_k^{(x)}} \right]^{-1} \left[ \sum_{k=1}^{N} z_k^{(x)\ast} z_k^{(y)} \right]$$

(8)

This estimate minimizes the weighted residual sum of squares (RSS)

$$RSS(c^2) = \sum_{k=1}^{N} \left[ z_k^{(y)} - z_k^{(y)\hat{h}} \right]^\ast V_k^{-1} \left[ z_k^{(y)} - z_k^{(y)\hat{h}} \right]$$

These results are complex generalizations of those presented by others (Ref. 21, section 1.5).

Analysis of the random effect. Statistical evaluation of the random subject effect is based on a likelihood ratio test, an essential component of which is the maximum likelihood estimate of $\sigma^2$. [A maximum likelihood estimate is the value of a population parameter that most likely would have produced the sample observations (see APPENDIX.)] The maximum likelihood estimate of $\sigma^2$ is obtained from a multivariate complex Gaussian distribution (12)

$$f[z^{(y)}] = \frac{1}{\pi^{m/2} \sigma^{2V}} \cdot \exp \left[ -\frac{1}{\sigma^{2}} \left[ z^{(y)} - z^{(y)\hat{h}} \right]^\ast V^{-1} \left[ z^{(y)} - z^{(y)\hat{h}} \right] \right]$$

For $N$ subjects, $l$, the $-2 \ln$ likelihood, is

$$l(h, \sigma^2, c^2) = \sum_{k=1}^{N} 2m \ln \pi + \ln |\sigma^{2V}|$$

$$+ \frac{1}{\sigma^{2}} \left[ z_k^{(y)} - z_k^{(y)\hat{h}} \right]^\ast V_k^{-1} \left[ z_k^{(y)} - z_k^{(y)\hat{h}} \right]$$

(9)

For a given value of $c^2$, the transfer function vector $h$ can be eliminated from the likelihood calculation by substituting Eq. 8 (to simplify, we set $h = \hat{h}$) into Eq. 9

$$l(\sigma^2, c^2) = 2m N \ln \pi + 2 \sum_{k=1}^{N} \ln |\sigma^{2V}| + \frac{2}{\sigma^{2}} RSS(c^2)$$

Because

$$\ln |\sigma^{2V}| = \ln (\sigma^{2m} |V_k|) = m \ln \pi + \ln |V_k|$$

the likelihood becomes

$$l(\sigma^2, c^2) = 2m N \ln \pi + 2m \ln \sigma^2$$

$$+ \frac{2}{\sigma^2} RSS(c^2)$$

(10)

Finally, differentiating Eq. 10 with respect to $\sigma^2$ and setting the derivative equal to zero gives the maximum likelihood estimate of $\sigma^2$

$$\hat{\sigma}^2 = \frac{1}{mN} RSS(c^2)$$

(11)

By substituting Eq. 11 into Eq. 10, the likelihood can now be written as a function of the single variable $c^2$, the ratio of the error variances ($\sigma_1^2/\sigma_2^2$)

$$l(c^2) = 2m N \ln (\pi \hat{\sigma}^2) + 2 \sum_{k=1}^{N} \ln |V_k| + 2mN$$

The determinant $|V_k|$, obtained by Cholesky factorization (13) of $V_k$, is

$$|V_k| = 1 + c^2 z_k^{(2)\ast} z_k^{(2)} = 1 + c^2 \sum_{n=1}^{m} |z_n^{(2)}|^2$$

Therefore, the likelihood is

$$l(c^2) = 2m N \ln (\pi \hat{\sigma}^2)$$

$$+ 2 \sum_{k=1}^{N} \ln \left[ 1 + c^2 z_k^{(2)\ast} z_k^{(2)} \right] + 2mN$$

(12)

the value of $c$ that minimizes Eq. 12 gives $\hat{\sigma}^2$ (Eq. 11) and the maximum likelihood estimates of the transfer functions (Eq. 7).

The example. Before we compare the $b$ estimates of the transfer function, we test whether the random effect for between-subject variability, $z_k^{(2)\ast} z_k^{(2)}$ in Eq. 6, improves the transfer function model. To do this, we first fit the fixed-effect (random effect absent) and mixed-effects (random effect present) models to all Fourier transforms from all subjects and obtain, for each model, values for $l$ and Akaike's information criterion (Table 3); the smaller Akaike's information criterion, the better the model. Then, using a likelihood ratio test, we evaluate the change in $-2 \ln$ likelihood, $\Delta l$, that occurs from adding the random subject effect. (APPENDIX discusses the test statistic used to evaluate $\Delta l$.) The large value of $\Delta l$ ($P < 0.0001$) signifies that the random subject effect does improve the transfer function model. The biological interpretation is that, for a given condition, transfer function estimates vary considerably among subjects.
Table 3. Transfer function estimates and contrasts

<table>
<thead>
<tr>
<th>Transfer Function Model</th>
<th>Without random effect $\gamma$</th>
<th>With random effect $\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline, $h_0$</td>
<td>$-972$</td>
<td>$-1.233$</td>
</tr>
<tr>
<td>Treatment, $h_1$</td>
<td>$-962$</td>
<td>$-1.221$</td>
</tr>
<tr>
<td>Test statistic (Eq. 14)*</td>
<td>5.44 ($P = 0.07$)</td>
<td>7.88 ($P = 0.02$)</td>
</tr>
<tr>
<td>$\Delta L$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akaike's information criterion</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Transfer function model is improved by including random effect $\gamma$ for between-subject variability: $\Delta L = -972 \sim (-1.233) = 261$, where $\Delta L$ is distributed as $\chi^2$ with 1 degree of freedom, $P < 0.0001$. *Test statistic for contrast between treatment (hypoxia-hypercapnia) and baseline transfer functions; under the null hypothesis that the contrast is 0, i.e., that the 2 transfer functions are identical, this test statistic is distributed as $\chi^2$ with 2 degrees of freedom.

**TRANSFER FUNCTION ANALYSIS**

In this section, we first review statistical contrasts and contrast coefficients; contrast coefficients are part of the test statistic used to compare transfer function estimates. Next, we detail the test statistic itself. Last, we compare $h_1$ to $h_0$, where $h_1$, $h_0$, $C$, the row vector of contrast coefficients in Eq. 14, must be

$$C = [-1 \ 1]$$

Compared with their respective values in the fixed-effect model, the mixed-effects estimates of $h_1$ and $h_0$ differ slightly. The larger test statistic for the $h_1 - h_0$ contrast reflects the greater power of the mixed-effects model.

The example. The transfer function estimates, obtained by using the mixed-effects model, are

$$\hat{h} = \left[ \begin{array}{c} \hat{h}_0 \\ \hat{h}_1 \end{array} \right] = \left[ \begin{array}{cc} 0.0047 & -i0.1446 \\ 0.0084 & -i0.1533 \end{array} \right]$$

where $\hat{h}_0$ and $\hat{h}_1$ estimate the transfer function during baseline and hypoxia-hypercapnia. To test whether hypoxia-hypercapnia alters the transfer function, we compare $h_1$ to $h_0$.

Test statistic. A statistical comparison between transfer function estimates requires the test statistic

$$2\hat{h}'C(C\hat{h}C' - 1)^{-1}C\hat{h}$$

where the matrix $C$, with row dimension $r$, contains the specific contrast coefficients, and where the covariance matrix for the transfer function estimates ($V_h$) is

$$V_h = \hat{\sigma}^2 \sum_{k=1}^N z_k(a)^* C V_h^{-1} z_k(a)^{-1}$$

The test statistic (Eq. 14) is similar to an $F$ statistic; under the null hypothesis that the difference (i.e., contrast) between two transfer functions is 0, it is distributed asymptotically as $\chi^2$ with 2$r$ degrees of freedom (see Ref. 36, p. 199). The factor of 2 in both the test statistic and its degrees of freedom is necessary because a transfer function has two components: its real and imaginary parts.

The explanation for the improved power of the mixed-effects model can be appreciated most easily by comparing the fixed-effect and mixed-effects models for subject $k$. When $b$ transfer functions are estimated for subject $k$, the fixed-effect model is

$$z_k^{(y\kappa)} = z_k^{(y)} h + z_k^{(a)}$$

The increase in power occurs because the fixed-effect error, $z_k^{(a)}$, is partitioned into two components in the mixed-effects model (Eq. 6): the within-subject (i.e., pure error) component $z_k^{(a)}$ and the random between-subject component $z_k^{(a)}$.

Confidence intervals for amplitude and phase. Confidence intervals for the amplitude and phase of the transfer function—or, by analogy, a transfer function effect model can be appreciated most easily by comparing the fixed-effect and mixed-effects models for subject $k$. When $b$ transfer functions are estimated for subject $k$, the fixed-effect model is

$$z_k^{(y\kappa)} = z_k^{(y\kappa)} h + z_k^{(a)}$$

The increase in power occurs because the fixed-effect error, $z_k^{(a)}$, is partitioned into two components in the mixed-effects model (Eq. 6): the within-subject (i.e., pure error) component $z_k^{(a)}$ and the random between-subject component $z_k^{(a)}$.

Confidence intervals for amplitude and phase. Confidence intervals for the amplitude and phase of the transfer function—or, by analogy, a transfer function contrast—can be calculated by using procedures derived by Groves and Hannan (14). The approximate 100(1 $- \alpha$)% confidence interval for the amplitude $A$ is

$$\hat{A} - z_{1/2}\hat{\sigma}_n/\sqrt{2} \leq A \leq \hat{A} + z_{1/2}\hat{\sigma}_n/\sqrt{2}$$

where $z_{1/2}$ is the critical value from the standard normal distribution and $\hat{\sigma}_n$ is the estimated SD of the transfer function $h$. The conservative 100(1 $- \alpha$)% confidence interval for the phase $\theta$ is

$$\hat{\theta} - \sin^{-1}[t_{1/2}\hat{\sigma}_n/\hat{A}] \leq \theta \leq \hat{\theta} + \sin^{-1}[t_{1/2}\hat{\sigma}_n/\hat{A}]$$

where $t_{1/2}$ is the critical value from a $t$-distribution with $2mN - 2$ degrees of freedom ($m$ is the number of frequencies within the frequency band and $N$ the number of subjects).
The example. Last, we calculate 95% confidence intervals for the amplitude and phase of the transfer function during baseline by using the inequalities in Eqs. 15 and 16 (Table 4). In addition to the estimates of amplitude and phase, $A$ and $\hat{\phi}$, respectively, and the estimated SD of the transfer function $\hat{s}_h$, we use the following information: $\alpha = 0.05$, $m = 15$, $N = 10$, $z_{0.025} = 1.96$, and $t_{0.025} = 1.65$ (298 degrees of freedom).

**DISCUSSION**

Previous approaches to transfer function estimation and analysis 1) assume erroneously that repeated observations on a subject are independent, 2) have limited statistical value and power, or 3) are restricted to use in single subjects rather than in an entire sample. The methodology we develop here is founded on a regression approach to time series analysis and corrects each of these deficiencies. By virtue of the random effect for between-subject variability in the transfer function estimates, our approach 1) can correctly model repeated observations on a subject, 2) has general relevance and increased statistical power, and 3) enables the simultaneous analysis of data from all subjects in a sample. Before discussing advantages and applications of this methodology, we review the use of a random subject effect in another statistical model developed for time series analysis.

Another random subject effect. Diggle and Al-Wasel (9) incorporate a random effect for between-subject variability in their model for the spectral density of an output time series measured during multiple experimental conditions. In our approach, it is the model for the transfer function between the input and output time series that has the random between-subject effect. When the transfer function is of primary experimental interest, our methodology is preferable.

Advantages of this new transfer function model. Our addition of the random effect $\gamma_k$ to the classic transfer function model has important theoretical and applied advantages. Three theoretical advantages pertain to the model itself. First, the new transfer function model (Eq. 6) accounts explicitly for biological differences that exist between subjects. Second, the model handles correctly the correlation among repeated observations on each subject. Third, the model is better quantitatively (see Table 3).

Three applied advantages relate to the use of this new transfer function model for the study of dynamic systems. First, the model provides for the simultaneous analysis of multiple subjects; the ability to properly analyze an entire sample is essential to draw conclusions about an underlying population. Second, the model provides for more accurate estimates of the mean and SD of the transfer function; the impact on the estimate of the SD, $\hat{s}_h$, is especially pronounced (see Table 4). Third, the model provides for an improved ability to detect differences between transfer functions (see Table 3).

Application to physiology. This improved statistical approach to transfer function estimation and analysis has broad relevance to the study of dynamic physiological systems. It can be used to estimate the mathematical relationship between virtually any pulsatile time series. In our animal model (Ref. 8), we estimate and analyze a transfer function at discrete frequencies: spontaneous heart rate and its first harmonic. Many physiological investigations, however, use random pacing to generate nearly continuous transfer function spectra. The methodology we present here accommodates this common situation also: transfer function analyses at multiple frequencies will be uncorrelated long as nonoverlapping frequency bands are used.

Summary. The transfer function methodology developed here furthers regression techniques to analyze the mathematical relationship between pulsatile time series. By correctly modeling repeated observations on each subject, this improved methodology can account for inherent between-subject variability in the transfer function estimates; this makes possible the simultaneous analysis of all subjects in a sample and improves the power of the transfer function model. This methodology can be used to estimate and analyze the dynamic properties of a variety of physiological systems.

**APPENDIX**

This APPENDIX shows how correlation decreases variability, details the test statistic used to evaluate the improvement gained by adding the random subject effect to the transfer function model, and reviews maximum likelihood estimation and statistical contrasts and contrast coefficients. Maximum likelihood estimation is used to obtain the transfer function estimates (Eq. 7). Contrast coefficients are a component of the test statistic (Eq. 14) used to compare transfer function estimates.

**Impact of Correlation on Variability**

Repeated observations on a subject are correlated (21) by virtue of that subject's particular biological makeup. When investigators fail to consider this correlation, as they usually...
do, they underestimate the true variability in the amplitude and phase estimates. To illustrate the impact of correlation on variability, imagine an investigation in a sample of \( n \) subjects in which the results are correlated, i.e., \( 0 \leq \rho \leq 1 \). The likelihood of a first measurement is reduced by a factor of \( \rho \). Given that the random variables \( B_1 \) and \( B_2 \) are considered jointly, then the distribution of the variable pair \((B_1, B_2)\) can be thought of as a joint probability density function (37), the bivariate normal distribution. For this bivariate normal distribution, \( \sigma_{21} \) is designated the SD of the conditional distribution of \( B_2 \) given that \( B_1 \) equals a specific value, depends on the correlation \( \rho \) between \( B_1 \) and \( B_2 \) (17):

\[
\sigma_{21} = \sigma_2 \sqrt{1 - \rho^2}
\]

where \( -1 \leq \rho \leq 1 \). Therefore, because repeated observations on a subject are correlated, i.e., \( \rho \neq 0 \), the SD of the variable measured during a second condition, given the value of the first measurement, is reduced by a factor of \( \sqrt{1 - \rho^2} \).

Maximum Likelihood Estimation

Sample observations are used to estimate population parameters: for example, the sample mean \( \bar{y} \) estimates the population mean \( \mu \). There are several techniques to estimate an unknown population parameter from a set of sample observations: the method of maximum likelihood is one (see Ref. 17).

To illustrate maximum likelihood estimation, the approach used in this methodology, imagine a sample of \( n \) independent observations, \( y_1, y_2, \ldots, y_n \), drawn from a population that is distributed normally with mean \( \mu \) and variance 1. The probability density function \( f \) describes the normal distribution of the population

\[
f(y) = \frac{1}{\sqrt{2\pi}} \cdot \exp \left[ -\frac{(y - \mu)^2}{2} \right]
\]

(A1)

Given the observations \( y_1, y_2, \ldots, y_n \) from this population, the likelihood \( L \) is considered to be a function of the population mean \( \mu \); it is the joint probability density function, the product of the \( n \) individual probability density functions (Eq. 17):

\[
L(\mu) = \prod_{i=1}^{n} f(y_i) = \left( \frac{1}{\sqrt{2\pi}} \right)^n \cdot \exp \left[ -\sum_{i=1}^{n} \frac{(y_i - \mu)^2}{2} \right]
\]

(A2)

in which the sample observations are substituted for the \( y_i \). The objective is to estimate \( \mu \) that is, to find the value of \( \mu \) that most likely would have produced the sample observations. This is achieved by solving for the value of \( \mu \) that maximizes the likelihood \( L \) (Eq. A2).

The value of \( \mu \) that maximizes \( L \) also maximizes the logarithm of \( L \). This is convenient: exponential terms are characteristic of many likelihood functions, and because products become sums, it is often simpler to maximize the logarithm of a likelihood rather than the likelihood itself. The logarithm \( l \) of the likelihood in Eq. A2 is

\[
l(\mu) = \ln L(\mu) = -\frac{n}{2} \ln (2\pi) - \sum_{i=1}^{n} \frac{(y_i - \mu)^2}{2}
\]

(A3)

The maximum likelihood estimate is the value of \( \mu \) for which

\[
\frac{dl(\mu)}{d\mu} = 0 \quad \text{and} \quad \frac{d^2l(\mu)}{d\mu^2} < 0
\]

Differentiating Eq. A3 with respect to \( \mu \) and setting the derivative equal to zero

\[
\frac{dl(\mu)}{d\mu} = \sum_{i=1}^{n} (y_i - \mu) = \sum_{i=1}^{n} y_i - n\mu = 0
\]

(A4)

yields the maximum likelihood estimate of \( \mu \)

\[
\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} y_i = \bar{y}
\]

That the estimate \( \hat{\mu} \) maximizes \( l \) (Eq. A3) can be confirmed by differentiating Eq. A4 with respect to \( \mu \)

\[
\frac{d^2l(\mu)}{d\mu^2} = -n
\]

and noting that \( -n < 0 \). In this example, the maximum likelihood estimate of the population mean \( \mu \) is the sample mean \( \bar{y} \).

Test Statistic Used to Evaluate Random Subject Effect

To test whether the random subject effect improves the transfer function model, we evaluate the corresponding change in \( l \), the \(-2 \ln \) likelihood (Eq. 12), by using a likelihood ratio test (see Analysis of the random effect). In this situation, the null hypothesis \( H_0 \) is that there is no between-subject variability in the transfer function estimates. This null hypothesis is written formally as

\[
H_0 : c^2 = 0
\]

where \( c^2 = \sigma^2_{21}/\sigma_2^2 \), the ratio of the error variances in the mixed-effects model. In Eq. 12, we minimize with respect to \( c^2 \); for any value of \( c^2 \), this constrains \( c^2 \) to be within the parameter space, or \( c^2 \geq 0 \). But \( H_0 : c^2 = 0 \) creates a problem for the likelihood ratio test: \( c^2 = 0 \) is on the edge of the parameter space. One solution is to differentiate the decision about the value of \( c^2 \), the maximum likelihood estimate of \( c^2 \), into two possibilities.

The first possibility is that \( c^2 = 0 \). In this case, the values of \( l \) for the fixed-effect (random effect absent) and mixed-effects (random effect present) models will be identical: therefore, \( \Delta l = 0 \), which can not be distributed as \( \chi^2 \). If \( \hat{c}^2 = 0 \), then the statistical decision is fail to reject \( H_0 \) and conclude that the data fail to demonstrate between-subject variability in the transfer function estimates.

The second possibility is that \( \hat{c}^2 > 0 \). In this case, \( H_0 : c^2 = 0 \) can be restated by moving it off the edge of the parameter space by an infinitesimal distance \( \epsilon \), where \( \epsilon > 0 \):

\[
H_0 : c^2 = \epsilon
\]

By doing this, the asymptotic distribution of \( \Delta l \) holds as \( \chi^2 \) with one degree of freedom (see Ref. 36, p. 124).

Statistical Contrasts

To illustrate contrasts and their coefficients, imagine an investigation of some biological variable during four experimental conditions: two control periods, during which the
values of the sample means are $\bar{y}_1$ and $\bar{y}_2$, and two intervention periods, during which the values of the sample means are $\bar{y}_3$ and $\bar{y}_4$. Suppose this investigation has two statistical objectives: 1) to compare the population means $\mu_1$ and $\mu_2$, and 2) to evaluate the average effect of the intervention on the population mean of the biological variable. Each objective is achieved by using a sample contrast $D$ (Eq. 13) to estimate a population contrast $D$.

The contrast used to make inferences about the population means $\mu_1$ and $\mu_2$ (objective 1) is

$$\hat{D}_1 = -\bar{y}_1 + \bar{y}_3$$

for which $\bar{y}_1$ is the arbitrary reference value. The coefficients for $D_1$ are

$$-1, 0, 1, 0$$

the second and fourth coefficients are 0 because $\bar{y}_2$ and $\bar{y}_4$ are absent from $D_1$.

The contrast used to make inferences about the average effect of the intervention on the population mean of the biological variable (objective 2) is

$$\hat{D}_2 = \left(-\frac{1}{2}\right)\bar{y}_1 + \left(\frac{1}{2}\right)\bar{y}_2 + \left(-\frac{1}{2}\right)\bar{y}_3 + \left(\frac{1}{2}\right)\bar{y}_4$$

The coefficients for $\hat{D}_2$ are

$$-\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}$$

We thank Dr. Mary Anne Farrell Epstein and the anonymous reviewers for their constructive comments and suggestions.

The development and analysis of the statistical methodology was funded by National Institute of General Medical Sciences Grant GM–38519 (to R. H. Jones). The physiological studies cited were funded by the Academic Enrichment Fund of the University of Colorado School of Medicine.

The authors are pleased to assist interested readers with the software that executes this methodology.

Address for reprint requests and software inquiries: D. Curran-Everett, Dept. of Pediatrics, B-195, Univ. of Colorado Health Sciences Center, 4200 East 9th Ave., Denver, CO 80262 (E-mail: dcurranever@castle.ucdenver.edu).

Received 11 December 1995; accepted in final form 5 August 1997.

REFERENCES