ORIGINAL ARTICLE

Normalization of organ bath contraction data for tissue specimen size: does one approach fit all?

Betul R. Erdogan[1] · Irem Karaomerlioglu[1] · Zeynep E. Yesilyurt[1] · Nihal Ozturk[1] · A. Elif Muderrisoglu[1] · Martin C. Michel[2] · Ebru Arioglu-Inan[1]

Abstract

Organ bath experiments are a key technology to assess contractility of smooth muscle. Despite efforts to standardize tissue specimen sizes, they vary to a certain degree. As it appears obvious that a larger piece of tissue should develop greater force, most investigators normalize contraction data for specimen size. However, they lack agreement which parameter should be used as denominator for normalization. A pre-planned analysis of data from a recent study was used to compare denominators used for normalization, i.e., weight, length, and cross-sectional area. To increase robustness, we compared force with denominator in correlation analysis and also coefficient of variation with different denominators. This was done concomitantly with urinary bladder strips and aortic rings and with multiple contractile stimuli. Our urinary bladder data show that normalization for strip weight yielded the tightest but still only moderate correlation (e.g., $r^2 = 0.3582$ for peak carbachol responses based on 188 strips). In aorta, correlations were even weaker (e.g., $r^2 = 0.0511$ for plateau phenylephrine responses normalized for weight based on 200 rings). Normalization for strip size is less effective in reducing data variability than previously assumed; the normalization denominator of choice must be identified separately for each preparation.

Keywords Normalization · Urinary bladder · Aorta · Weight · Length · Cross-sectional area

Introduction

Organ bath experiments are a widely used technology to study smooth muscle responsiveness to neurotransmitters, paracrine factors, hormones, and xenobiotics. It is generally assumed that a larger specimen can intrinsically develop a greater force of contraction than a smaller one; therefore, investigators typically attempt to generate tissue specimens of comparable size. However, despite these efforts, some variability of sizes exists. Differences in force of contraction between specimens can become important when comparing groups of animals or humans, e.g., healthy and diseased ones. Therefore, many investigators attempt to normalize observed force of contraction for specimen size.

While the agreement to normalize force of contraction based upon size appears widely accepted, investigators used different denominators for normalization. Typical choices, e.g., in blood vessels include strip weight (Kershaw et al. 2004; Kunert et al. 2010; Yang et al. 2018), length (Fujishige et al. 2002; Török et al. 2016) or cross-sectional area (CSA) (Wyse 1980; Abebe et al. 1990; Cameron and Cotter 1992; Weber et al. 1996; Pieper et al. 1997; Conklin and Boor 1998; Ozcelikay et al. 2000; Kristek et al. 2009). While CSA appears to be the most popular choice as denominator for normalization of artery contraction, this is not necessarily the case in other tissues. For instance, in the urinary bladder normalization for weight (Paro et al. 1990; Kories et al. 2003; Stevens et al. 2006), length (Schneider et al. 2005a) and CSA (Braverman and Ruggieri Sr. 2003) appear to be similarly popular. Surprisingly, most investigators do not justify or otherwise discuss their choice of denominator.

Only few studies have compared denominators used for normalization (Wyse 1980; Schneider et al. 2005b; Jin et al.

✉ Martin C. Michel
  marmiche@uni-mainz.de

1 Department of Pharmacology, Faculty of Pharmacy, Ankara University, Ankara, Turkey

2 Department of Pharmacology, Johannes Gutenberg University, Mainz, Germany

2018). Typically, these were based on limited sample sizes, lacked a replication cohort, and were based on a single tissue. The latter is relevant because it is possible that a normalization approach that is useful in one tissue may be less useful in another. Moreover, the usefulness of a normalization denominator may at least hypothetically not be the same across contractile stimuli or parameters describing contraction. Therefore, we have explored whether normalization by weight, length, or CSA yields more relevant results. We have done so in two tissues (rat urinary bladder and aorta); investigated different receptor-mediated and receptor-independent contractile stimuli (carbachol, phenylephrine, KCl), peak and plateau contractile responses, and agonist potency; and applied multiple approaches to test for effectiveness of normalization (correlation between force and denominator and coefficient of variation (CV) between types of normalized data). To enhance robustness, these analyses were done in two batches of animals and in the pooled data. All of these analyses were done as part of a pre-planned secondary analysis of a recently published study (Yesilyurt et al. 2019).

Material and methods

Animals and treatments

Our analysis is based on samples from a study to explore effects of experimental diabetes and its treatment with the sodium-glucose co-transporter-2 inhibitor dapagliflozin on the heart, aorta, and urinary bladder (Yesilyurt et al. 2019). Male Sprague Dawley rats (5 weeks old) were obtained from Bilkent University Genetics and Biotechnology Research Center (Ankara, Turkey) and housed under 12:12-h lighting conditions with free access to standard chow (Purina Rat Chow (5% fat); Optima AS, Bolu, Turkey) or high-fat diet (HFD; OpenSource diet, D12492 (35% fat); Arden Research & Experiment, Ankara, Turkey) and tap water.

The underlying study was performed in two batches of 6 and 10 animals per group, respectively, studied 6 months apart for logistic reasons (Yesilyurt et al. 2019). The animals from each batch were divided into four groups (based on randomization for the second batch): group I were healthy controls, group II was diabetic as induced by HFD plus low-dose streptozotocin (STZ; 25 mg/kg intraperitoneal, dissolved in citrate buffer at pH 4.5, injected when 10–11 weeks old), group III were healthy rats treated with dapagliflozin (1 mg/kg/day by oral injection), and group IV were treated diabetic rats (HFD plus low-dose STZ plus dapagliflozin). Animals in group II and group IV received a second or third STZ injection if the blood glucose levels were < 200 and < 140 mg/dl in batches 1 and 2, respectively. Treatment with dapagliflozin was started in groups III and IV at the age of 18–21 weeks.

Experimental procedure

Animals were killed by exsanguination under anesthesia with inhalation of 2% isoflurane (first batch) or diethyl ether (second batch) at a time point corresponding to 18–26 weeks after STZ injection (i.e., 13–15 weeks and 12–13 weeks after start of dapagliflozin treatment in batch 1 and 2, respectively). The urinary bladder and aorta were excised, and adjacent adipose and soft connective tissue were removed. Based on previous validation experiments (Schneider et al. 2011), strips were stored in ice-cold Krebs-Henseleit buffer for up to 2 h prior to use in the organ bath.

Organ bath experiments were performed as described previously for bladder (Michel 2014) with minor modifications. Briefly, the upper most dome and the lower trigone area were removed, and the remaining body of the bladder was cut into four longitudinal strips of approximately 1–2 mm width. The strips were mounted in a 10-ml organ bath in Krebs-Henseleit buffer (118.5 mM NaCl, 4.7 mM KCl, 1.2 mM $MgSO_4$, 2.5 mM $CaCl_2$, 1.2 mM KH_2PO_4, 25 mM $NaHCO_3$, and 5.6 mM glucose continually gassed with 95% O_2/5% CO_2 to maintain a pH of 7.4 at 37 °C) under a resting tension of 10 mN. They were allowed 75 min of equilibration including washes with fresh buffer every 15 min and re-adjustment of resting tension after each wash. Each strip was challenged twice with 50 mM KCl (maintaining iso-osmolarity by reducing NaCl concentration from 116.8 to 68.5 mM) with 60 min rest between challenges; the peak response to the second KCl addition was used to describe receptor-independent contraction. After washing and an additional 45 min equilibration, a carbachol concentration-response curve was generated by adding cumulative concentrations of the muscarinic agonist carbachol (10 nM–300 μM) in half-logarithmical steps, and peak tension was measured for each concentration. After the highest carbachol concentration, strips were washed and allowed another 45 min of equilibration. Thereafter, 1 μM carbachol was added. When steady-state tension was reached, a cumulative concentration-response curve for a β-adrenoceptor agonist was generated.

Organ bath experiments with aorta were performed as previously described (Ozcelikay et al. 2000) with minor modifications. Thoracic aorta rings were mounted at a resting tension of 19.6 mN in 10 ml organ bath filled with Krebs-Henseleit buffer (120 mM NaCl; 4.8 mM KCl; 1.25 mM $CaCl_2$; 1.25 mM $MgSO_4$; 1.2 mM KH_2PO_4; 25 mM $NaHCO_3$, and 10 mM glucose continually gassed with 95% O_2/5% CO_2 to maintain a pH of 7.4 at 37 °C). Aortic rings were equilibrated for 1 h with washes with fresh buffer every 15 min. Thereafter, rings were challenged with 10 μM phenylephrine; presence of functional endothelium was checked for relaxation response to 10 μM acetylcholine once the phenylephrine response had reached a plateau; this contraction level immediately prior to addition of acetylcholine was defined as plateau response. In batch 1,

preparations were subsequently washed and exposed to 10 nM–30 µM phenylephrine (in half-logarithmic steps); peak contraction responses were measured and used to construct a concentration-response curve. The calculated maximum response calculated from this curve was regarded as peak phenylephrine response. At the end of each experiment, the weight and the length of the bladder strips and aorta rings were measured, and CSA was calculated as (weight/(length × 1.05)).

Chemicals

Dapagliflozin tablets (Forziga™) were obtained from Astra Zeneca (Ankara, Turkey), crushed in a mortar, and suspended in distilled water. The materials for Krebs solution, STZ, carbachol, and phenylephrine were obtained from Sigma Aldrich (Ankara, Turkey). Isoflurane and diethyl ether were from Adeka (Samsun, Turkey) and from Merck (Ankara, Turkey), respectively.

Data analysis and quality measures

All analyses in the present manuscript had been pre-specified prior to obtaining experimental data except otherwise noted as a pre-planned secondary analysis of a study reported elsewhere (Yesilyurt et al. 2019). Sample size was defined as using all bladder strips and aorta rings from the underlying study. In the underlying study, contractile response to KCl and carbachol in the bladder and to phenylephrine in the aorta were similar in all four treatment groups (Yesilyurt et al. 2019). Therefore, we used strips from all four groups for the present analyses. For reasons of the underlying study, peak contraction responses to phenylephrine in aorta were only measured in batch 1, whereas plateau responses were measured in both batches. As our approach implied that variability in force of contraction derives at least in part from the dimensions of a tissue strip, each strip was considered as a biological replicate. Our original protocol specified to analyze each batch in isolation. To further increase robustness, a post-hoc decision was made to additionally analyze pooled data from both batches.

Randomization of treatment sequence was applied to batch 2. All experiments and analyses of both batches were performed by investigators blinded for group allocation; blinding included not only concealment of group allocation, but also of blood glucose and body weight because the latter may have indirectly unblinded the assessment. Unblinding was performed after all experimental data of a given batch had been analyzed. No data points or experiments ("outliers") were removed from the analysis.

Based on previous studies (Schneider et al. 2004b, 2005b; Frazier et al. 2007; Michel 2014), we were aware that contractile responses to carbachol in the bladder typically reach a maximum at 30 µM of the agonist with smaller contractions at higher concentrations, i.e., exhibit a bell-shaped curve.

Because the nature of the declining part of the concentration-response curve is unclear, we pre-specified to analyze only responses for 10 nM–30 µM in the control group. However, we also tested 100 and 300 µM carbachol in all groups to allow detection of possible right-shifts of the curve upon treatment. As we observed a decline of contractions at 100 and 300 µM carbachol in every single rat of all treatment groups, we also restricted curve fitting in the treated groups to the 10 nM–30 µM range, which is in line with our previous studies (Schneider et al. 2004b, 2005b; Frazier et al. 2007; Michel 2014). Plateau responses to carbachol were based on a single concentration of 1 µM only. In aorta plateau, responses to a single concentration of 10 µM phenylephrine were obtained immediately prior to addition of acetylcholine.

While only peak responses to KCl were analyzed, we measured both peak and plateau responses to carbachol and phenylephrine. They were expressed as raw data after normalization for weight, length, and CSA of the specimen. Data were analyzed in two ways. Firstly, we correlated observed force of contraction with strip length, weight, and CSA; the resulting Pearson's r^2 values were taken as indicators of strength of association. We assumed that the indicator of strip size exhibiting the strongest association was most suitable for normalization. Second, in a post-hoc analysis, we additionally determined coefficient of variation (CV) for each parameter. This was done because on a conceptual level normalization is done to reduce overall variability by taking one source (strip size) out of the equation. Correlations of agonist potency with indicators of strip size have been included as negative controls because there are no logical reasons why potency should be related to strip size.

Due to the exploratory nature of our analyses, no hypothesis-testing statistical analyses were performed; reported p values of the correlation analyses are descriptive only. Rather, we looked for consistency between both batches of animals. Data are reported as means ± SD. All curve fitting and statistical analyses were performed with Prism (version 8.2, GraphPad, La Jolla, CA, USA).

Results

Bladder

We obtained 95 strips from 24 rats of batch 1 and 93 strips from 39 rats of batch 2. In the pooled data of both batches, mean strip weight was 18.96 ± 6.87 mg (range 3.50–42.70), length 16.63 ± 4.23 mm (7.00–28.00), and CSA 1.13 ± 0.43 mg/(mm·1.05) (0.30–2.61). Mean peak contraction responses to 50 mM KCl were 4.75 ± 1.52 mN, which corresponded to 0.27 ± 0.12 mN/mg, 0.30 ± 0.12 mN/mm, and 4.62 ± 1.85 mN/(mg/(mm·1.05)). Corresponding values for peak carbachol responses derived from concentration-response curves were 9.30 ± 2.44 mN, 0.53 ± 0.20 mN/mg,

0.60 ± 0.23 mN/mm, and 9.01 ± 2.96 mN/(mg/(mm·1.05)). For plateau carbachol responses, they were 2.78 ± 0.89 mN, 0.16 ± 0.07 mN/mg, 0.17 ± 0.06 mN/mm, and 2.74 ± 1.08 mN/(mg/(mm·1.05)). Mean pEC_{50} of carbachol was 6.04 ± 0.31 (5.09–7.46).

Figure 1 shows the correlation of peak responses to KCl and carbachol and plateau responses to carbachol with strip weight, length, and CSA. The square of the coefficient of correlation (r^2) for each contractile stimulus in each batch and in the pooled data is shown in Table 1. The strength of correlation was only moderate at best for any combination of contractile stimulus and normalization parameter, and this was consistent across both batches and the pooled data. Across all three stimulation conditions, correlations were relatively strongest for strip weight ($r^2 = 0.21–0.41$), somewhat weaker

for strip CSA, and weakest for strip length. In some cases, the strength of correlation with strip length had a descriptive p value larger than 0.05. Similar to force, the potency (pEC_{50}) of carbachol exhibited also only poor if any correlation with strip weight, length, or CSA (Table 1). Similar conclusions were reached when each of the four treatment groups was considered in isolation (Online Supplemental Table 1). The CV of contraction data was comparable for raw data and those normalized for weight, length, or CSA (Table 2).

Aorta

We obtained 96 aortic rings from 24 rats of batch 1 and 104 rings from 39 rats of batch 2. Mean ring weight, length, and CSA in pooled data of both batches were 7.01 ± 1.61 mg
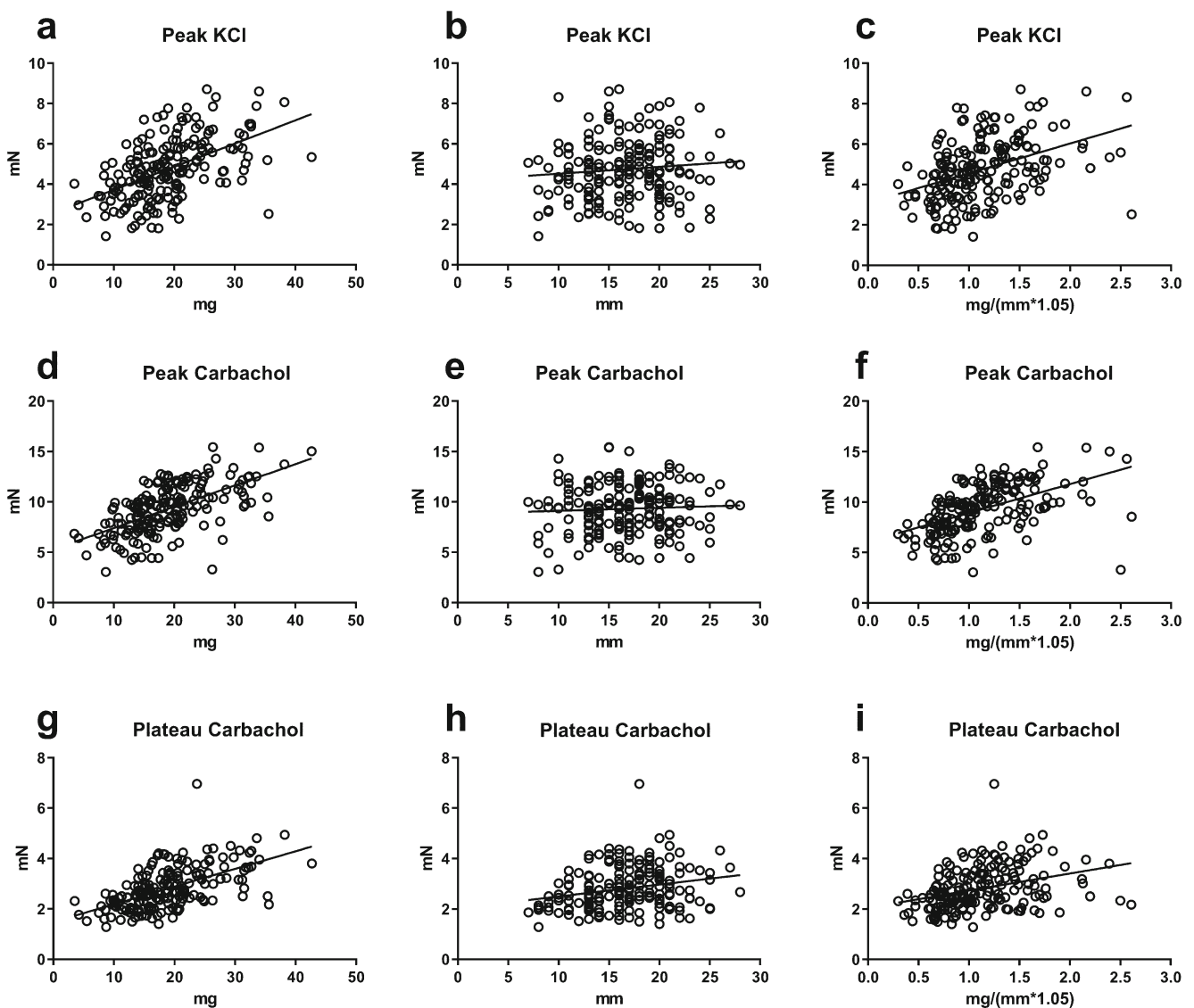

Fig. 1 Correlation of peak contractile response of urinary bladder strips to 50 mM KCl (panels **a**–**c**) and carbachol (panels **d**–**f**) and plateau response to 1 μM carbachol (panels **g**–**i**) to strip weight (panels **a**, **d**, and **g**), length (panels **b**, **e**, and **h**) and cross-sectional area (CSA; panels **c**, **f**, and **i**) in the pooled data of both batches (figures were generated using GraphPad Prism, version 8.2). For quantitative data see Table 1

Table 1 Correlation between potency of carbachol of urinary bladder strips in batch 1 and correlation between contractile responses of urinary bladder strips to peak responses to KCl and carbachol and plateau responses to carbachol with strip weight, length and cross-sectional area (CSA) in batches 1 and 2 and pooled data. Shown are squared coefficient of correlation (r^2) and corresponding descriptive p values. A graphical representation of the correlations based on the pooled data is shown in Fig. 1

	Force vs. weight (mg)		Force vs. length (mm)		Force vs. CSA (weight/(length · 1.05))	
	r^2	p value	r^2	p value	r^2	p value
Batch 1						
KCl peak	0.2908	< 0.0001	0.0026	0.6237	0.2162	< 0.0001
Carbachol peak E_{max}	0.4022	< 0.0001	0.0343	0.0790	0.1941	< 0.0001
Carbachol peak pEC_{50}	0.0004	0.8560	0.0596	0.0197	0.0403	0.0565
Carbachol plateau	0.2099	< 0.0001	0.0651	0.0147	0.0321	0.0891
Batch 2						
KCl peak	0.2907	< 0.0001	0.0149	0.2433	0.1289	0.0004
Carbachol peak	0.4065	< 0.0001	0.0098	0.3469	0.3556	< 0.0001
Carbachol plateau	0.2895	< 0.0001	0.0335	0.0825	0.0849	0.0051
Pooled data						
KCl peak	0.2697	< 0.0001	0.0091	0.1929	0.1751	< 0.0001
Carbachol peak	0.3582	< 0.0001	0.0026	0.4929	0.2615	< 0.0001
Carbachol plateau	0.3306	< 0.0001	0.0556	0.0014	0.1160	< 0.0001

(3.70–12.30), 5.76 ± 0.92 mm (3.00–8.00), and 1.16 ± 0.23 mg/(mm·1.05) (0.76–2.69). Based on the needs of the underlying study, peak tension responses based on concentration-response curves were only measured in batch 1. Mean peak responses to phenylephrine were 3.11 ± 0.94 mN, 0.40 ± 0.12 mN/mg, 0.49 ± 0.16 mN/mm, and 2.65 ± 0.78 mN/(mg/(mm·1.05)). Plateau responses to a submaximal concentration of phenylephrine were measured in both batches. The mean plateau response to phenylephrine was 2.83 ± 0.73 mN, which corresponded to 0.42 ± 0.12 mN/mg, 0.50 ± 0.14 mN/mm, and 2.49 ± 0.68 mN/(mg/(mm1.05))·. Mean pEC_{50} of phenylephrine in batch 1 was 7.38 ± 0.61 (5.85–8.53).

Figure 2 shows the correlation of peak and plateau responses to phenylephrine with ring weight, length, and CSA. The square of the coefficient of correlation for peak and plateau response to phenylephrine in each batch and in the pooled data is given in Table 3. The strength of correlation between force and either ring weight, length, or CSA was very weak to non-existent ($r^2 < 0.06$ and descriptive p value > 0.05 in all cases). Similar to force, potency of phenylephrine exhibited also only poor if any correlation with ring weight, length, or CSA (Table 3). Similar conclusions were reached when each of the four treatment groups was considered in isolation (Online Supplemental Table 2). The CV of contraction data was comparable for raw data and those normalized for weight, length, or CSA (Table 2).

Discussion

It appears obvious that force of contraction of tissue specimens should depend on their size. Absolute force of contraction may be irrelevant for some research questions, for instance when the effect of multiple concentrations of an inhibitor is tested during repeated concentration-response curves within a single tissue specimen (Schneider et al. 2004a; Klausner et al. 2009; Sand and Michel 2014) or when force for the stimulator of interest is normalized to an internal control agonist (Meini et al. 1998; Conklin et al. 2004, 2009; Michel 2014). If force is compared across groups of strips, measured values should somehow be normalized for size of tissue specimen. However, there is no uniform agreement how this should be done with some investigators using strip length, weight, or CSA as denominator (see the "Introduction" section).

Critique of methods

We have used several design elements to increase robustness of our findings. Firstly, all analyses expect specifically noted otherwise had been pre-specified before data were obtained. Second, we had also pre-specified sample size as all available strips from the underlying study. In this regard, we considered each specimen to be a biological replicate as the possible relationship between size and force should depend on the size and contraction response of an individual specimen. Third, we have explored multiple parameters of contractility, i.e., peak and plateau contraction forces and agonist potency. Fourth, we have used three different contractile stimuli, the receptor agonists carbachol and phenylephrine and the receptor-independent stimulus KCl. For the agonists, we looked at peak and plateau contraction responses. Fifth, the underlying study had applied randomization (for batch 2) and blinding for group allocation for all experimental steps including data analysis,

Table 2 Coefficient of variation (CV) of raw data and those normalized for strip/ring weight, length, and cross-sectional area (CSA) for various contractile stimuli in rat urinary bladder and aorta

	Raw data	Force/mg	Force/mm	Force/CSA
Urinary bladder				
Batch 1				
KCl peak	33.73	32.25	41.71	33.21
Carbachol peak	27.79	27.36	34.46	29.70
Carbachol plateau	34.68	33.30	36.19	35.50
Batch 2				
KCl peak	30.19	43.71	38.49	44.25
Carbachol peak	24.63	37.68	39.55	34.66
Carbachol plateau	26.90	43.63	34.08	43.10
Pooled data				
KCl peak	32.04	42.23	40.54	40.03
Carbachol peak	26.21	37.19	38.76	32.90
Carbachol plateau	31.98	40.76	35.18	39.45
Aorta				
Batch 1				
Phenylephrine peak	30.09	29.56	31.98	29.47
Phenylephrine plateau	29.14	30.05	30.45	29.96
Batch 2				
Phenylephrine plateau	20.16	24.01	23.22	24.43
Pooled data				
Phenylephrine plateau	25.78	27.86	27.02	27.30

thereby minimizing observer bias. Sixth, we have analyzed both batches of the underlying study separately to check for consistency of our findings. Each batch was at least as large, mostly larger, than any previous reported similar analysis. Finally, and in contrast to previously reported studies of a similar type (Wyse 1980; Schneider et al. 2005b; Jin et al. 2018), we have based our analyses in two tissues, rat urinary bladder, and aorta.

The present study analysis was designed to be exploratory, i.e., did not test a pre-specified statistical null hypothesis. Therefore, any *p* value reported here should be considered as descriptive and not as hypothesis-testing. Rather looking separately at two batches of experiments was used to reduce the probability of a chance finding. The consistent degree of correlation (or lack of it) across batches confirms the validity of this approach.

Our specimens came from a study with four treatment arms. Our analysis is based on the pooled specimens from all four arms. While it is hypothetically possible that the relationship between specimen size and force is modified by disease (diabetes) and/or treatment (dapagliflozin), we consider this to be a very unlikely source of variability because contractile responses were very similar across the four arms (Yesilyurt et al. 2019). As a result, the number of pairs in our correlation analysis markedly exceeds that of any previously reported comparison by far.

Choice of denominator for normalization

The importance of critically evaluating denominators chosen for normalization of contraction data is illustrated by reports where strip weight and CSA but not length differed between groups (Longhurst et al. 1990) or where detection of a difference between groups depended on the denominator being applied for normalization (Schneider et al. 2005b; Jin et al. 2018).

In our earlier rat bladder work, we reported based on post-hoc analysis that peak contraction responses to carbachol were correlated with strip weight ($r^2 = 0.664$), but this was based on 24 strips only (Kories et al. 2003). In a later study, we found that carbachol responses exhibited a tighter association with strip length than weight ($r^2 = 0.362$ vs. 0.130) (Schneider et al. 2005a) and used normalization by length from that study onward. While this was based on a somewhat larger sample size ($n = 49$), it still represented a post-hoc analysis. Similar analyses by others for blood vessels have also relied on limited sample sizes ($n = 9$–20) (Wyse 1980; Jin et al. 2018). We have now repeated such analyses using both bladder and aorta, a pre-specified design, and considerably larger sample sizes ($n = 93$–104 strips per batch). We consider this larger sample size and the use of two batches to be important because our previous analyses found an r^2 for the correlation with weight of either 0.664 (Kories et al. 2003) or 0.130 (Schneider et al.
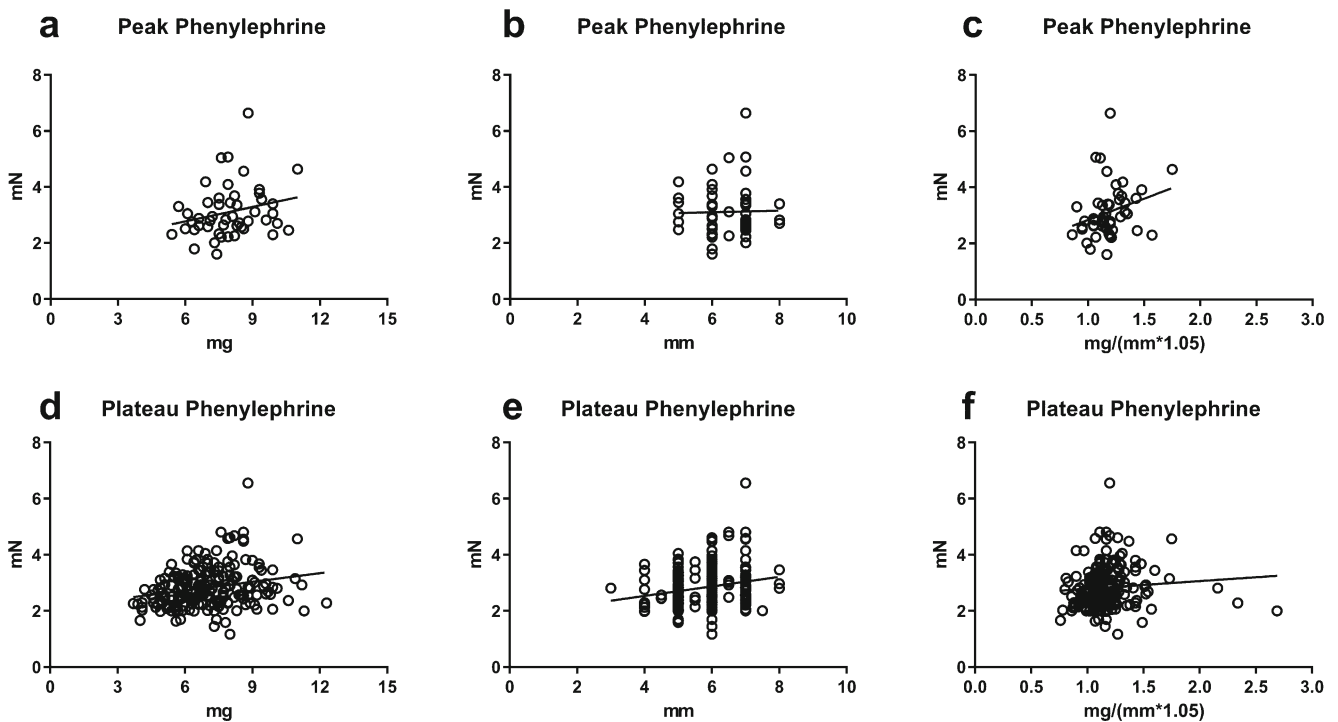
Fig. 2 Correlation of peak contractile response of aortic rings to phenylephrine (concentration-response curve; panels **a**–**c**) and plateau response to 10 μM phenylephrine (panels **d**–**f**) to ring weight (panels **a** and **d**), length (panels **b** and **e**), and cross-sectional area (CSA; panels **c** and **f**) in the pooled data of both batches (figures were generated using GraphPad Prism, version 8.2). For quantitative data, see Table 3

2005a), highlighting the inherent danger of inference based on small sample size (Halsey et al. 2015).

We found only weak to moderate correlations between specimen size and force of contraction, irrespective of tissue or contractile stimulus, and this was consistent across batches. Some associations, particularly those in the aorta, were so weak that their descriptive *p* value remained high even with the very large sample size of the pooled batches. Similarly, applying normalization did not reduce variability assessed as CV but, if anything, slightly increased it. While we cannot explain the discrepancy

between an at least moderate correlation but a similar CV in the bladder, both approaches suggest that normalization for strip size is not very effective in reducing variability.

Despite appearing counter-intuitive, these data suggest that specimen size contributes in a limited way only to observed inter-sample variability in contractile responses in the bladder. This is qualitatively in line with our previous analysis based on much smaller sample sizes (Kories et al. 2003; Schneider et al. 2005a). While our data do not demonstrate a lower CV upon normalization, they show a moderate correlation

Table 3 Correlation between contractile responses of aorta rings to peak responses to phenylephrine and correlation between potency of phenylephrine in batch 1 and plateau responses to phenylephrine in batch 1, batch 2 and pooled data with rings weight, length, and cross-sectional area (CSA). Shown are squared coefficient of correlation (r^2) and corresponding descriptive p values. A graphical representation of the correlations based on the pooled data is shown in Fig. 2

	Force vs. weight (mg)		Force vs. length (mm)		Force vs. CSA (weight/(length × 1.05))	
	r^2	*p* value	r^2	*p* value	r^2	*p* value
Batch 1						
Phenylephrine peak E_{max}	0.0564	0.1005	0.0006	0.8719	0.0735	0.0596
Phenylephrine pEC_{50}	0.0105	0.4842	0.0083	0.5326	0.0012	0.8154
Phenylephrine plateau	0.0295	0.0942	0.0053	0.4795	0.0258	0.1177
Batch 2						
Phenylephrine plateau	0.0147	0.2195	0.0418	0.0375	0.0000	0.9706
Pooled data						
Phenylephrine plateau	0.0511	0.0013	0.0466	0.0021	0.0074	0.2246

between strip weight and contractile responses. The latter would suggest that normalization for strip weight is beneficial in rat urinary bladder strips in reducing inter-strip variability. Variability is one of the determinants of calculations of statistical power, and a reduction of variability by about 50% has a similar impact on power as a quadrupling of sample sizes. Therefore, normalization of bladder contraction data for strip size could have an impact on required sample sizes and assist investigators in adhering to the 3R principle ("replace, reduce, refine") of ethical use of experimental animals (Kilkenny et al. 2010). However, our CV data do not support this conclusion.

Our data in aorta partly agree and partly disagree with those in the bladder. They agree that normalization for size has limited impact at best. However, they disagree because we did not find evidence for a contribution of any indicator of ring size on observed variability in contractile response. This contrasts earlier reports by others (Wyse 1980; Jin et al. 2018) who reported beneficial effects of normalization for blood vessels, specifically for the use of CSA as denominator. However, these previous studies were based on small sample sizes of 9–20. Our own experience in the bladder (Kories et al. 2003; Schneider et al. 2005a) suggests that small sample sizes may be misleading, particularly if publication bias is taken into consideration.

In conclusion, we found that normalization of organ contraction data for strip size has only limited impact on observed variability. Apparently, factors other than measured strip size (be in length, weight, or CSA) explain the majority of inter-strip differences in contractile responses. While these other factors, possibly including imprecise measurements or connective tissue content, remain unknown, our data demonstrate that normalization based on widely used indicators of strip size is helpful to a limited extent at best. Nonetheless, we recommend that normalization efforts should be continued to reduce variability, even if to a minor degree only, as this improves statistical power and thereby leads to a more ethical use of experimental animals (Kilkenny et al. 2010). In this regard, it is surprising that normalization is applied by most investigators, but only very few have reported on the effectiveness of their normalization approach. Our data also suggest that exploration of suitable denominators for normalization may need to be tested specifically for each tissue. Lessons on choice of denominator for normalization of experimental data are likely to also be applicable to experimental techniques other than organ bath studies (Michel-Reher and Michel 2015). On a more practical level, we make two main recommendations:

– Authors should carefully check (and document in their reports) which normalization parameter is most appropriate in the model under investigation. They should not necessarily adopt a certain denominator for normalization just because others have done so.

– Normalization of responses using a receptor-independent denominator such as responses to KCl is a double-edged sword. On the one hand, this approach allows to explore possible alterations of receptor responsiveness without assumptions on choice of adjustment for strip size. On the other hand, this approach will not be helpful to explore downstream alterations of receptor responsiveness, i.e., those shared by response to KCl, for instance those in calcium handling by smooth muscle. If KCl is used as denominator for normalization, quantitative KCl responses should be reported for each group so that readers can see whether or to which degree alterations in receptor responses may in part be driven by those to KCl.

Compliance with ethical standards The study protocol had been approved by the animal welfare committee of Ankara University (permit 2015-4-82, 2017-10-92) and was in line with NIH Guidelines for Care and Use of Laboratory Animals.

Conflict of interest The authors report no conflict of interest related to this paper.

References

Abebe W, Harris KH, MacLeod KM (1990) Enhanced contractile responses of arteries from diabetic rats to α_1-adrenoceptor stimulation in the absence and presence of extracellular calcium. J Cardiovasc Pharmacol 16:239–248

Braverman AS, Ruggieri MR Sr (2003) Hypertrophy changes the muscarinic receptor subtype mediating bladder contraction from M_3 toward M_2. Am J Phys 285:R701–R708

Cameron NE, Cotter MA (1992) Impaired contraction and relaxation in aorta from streptozotocin-diabetic rats: role of polyol pathway. Diabetologia 35:1011–1019

Conklin DJ, Boor PJ (1998) Allylamine cardiovascular toxicity: evidence for aberrant vasoreactivity in rats. Toxicol Appl Pharmacol 148:245–251

Conklin DJ, Cowley HR, Wiechmann RJ, Johnson GH, Trent MB, Boor PJ (2004) Vasoactive effects of methylamine in isolated human blood vessels: role of semicarbazide-sensitive amine oxidase, formaldehyde, and hydrogen peroxide. Am J Phys 286:H667–H676

Conklin DJ, Haberzettl P, Prough RA, Bhatnagar A (2009) Glutathione-S-transferase P protects against endothelial dysfunction induced by exposure to tobacco smoke. Am J Phys 296:H1586–H1597

Frazier EP, Braverman AS, Peters SLM, Michel MC, Ruggieri MR Sr (2007) Does phospholipase C mediate muscarinic receptor-induced rat urinary bladder contraction? J Pharmacol Exp Ther 322:998–1002

Fujishige A, Takahashi K, Tsuchiya T (2002) Altered mechanical properties in smooth muscle of mice with a mutated calponin locus. Zool Sci 19:167–174

Halsey LG, Curran-Everett D, Vowler SL, Drummond GB (2015) The fickle P value generates irreproducible results. Nat Med 12:179–185

Jin L, Lipinski A, Conklin DJ (2018) A simple method for normalization of aortic contractility. J Vasc Res 55:177–186

Kershaw JDB, Misfield M, Sievers H-H, Yacoub MH, Chester AH (2004) Specific regional and directional contractile responses of aortic cup tissue. J Heart Valve Dis 13:798.803

Kilkenny C, Browne W, Cuthill IC, Emerson M, Altman DG (2010) Animal research: reporting in vivo experiments: the ARRIVE guidelines. Br J Pharmacol 160:1577–1579

Klausner AP, Rourke KF, Miner AS, Ratz PH (2009) Potentiation of carbachol-induced detrusor smooth muscle contractions by ß-adrenoceptor activation. Eur J Pharmacol 606:191–198

Kories C, Czyborra C, Fetscher C, Schneider T, Krege S, Michel MC (2003) Gender comparison of muscarinic receptor expression and function in rat and human urinary bladder: differential regulation of M_2 and M3? Naunyn Schmiedeberg's Arch Pharmacol 367:524–531

Kristek F, Cacanyiova S, Gerova M (2009) Hypotrophic effect of long-term neuronal NO-synthase inhibition on heart and conduit arteries of the Wistar rats. J Physiol Pharmacol 60:21–27

Kunert MP, Dwinell MR, Lombard JH (2010) Vascular responses in aortic rings of a consomic rat panel derived from the fawn hooded hypertensive strain. Physiol Genomics 42A:244–258

Longhurst PA, Kang J, Wein AJ, Levin RM (1990) Length-tension relationship of urinary bladder strips from streptozotocin-diabetic rats. Pharmacology 40:110–121

Meini S, Lecci A, Cucchi P, Catalioto RM, Criscuoli M, Maggi CA (1998) Inflammation modifies the role of cyclooxygenases in the contractile responses of the rat detrusor smooth muscle to kinin agonists. J Pharmacol Exp Ther 287:137–143

Michel MC (2014) Do ß-adrenoceptor agonists induce homologous or heterologous desensitization in rat urinary bladder? Naunyn Schmiedeberg's Arch Pharmacol 387:215–224

Michel-Reher MB, Michel MC (2015) Regulation of GAPDH expression by treatment with the ß-adrenoceptor agonist isoprenaline - is GAPDH a suitable loading control in immunoblot experiments? Naunyn Schmiedeberg's Arch Pharmacol 388:1119–1120

Ozcelikay AT, Tay A, Güner S, Tasyaran V, Yildizoglu-Ari N, Dincer UD, Altan VM (2000) Reversal effects of L-arginine treatment on blood pressure and vascular responsiveness of streptozotocin-diabetic rats. Pharmacol Res 41:201–209

Paro M, Italiano G, Travagli RA, Petrelli L, Zanoni R, Prosdocimi M, Fiori MG (1990) Cystometric changes in alloxan diabetic rat: evidence for functional and structural correlates of diabetic autonomic neuropathy. J Auton Nerv Syst 30:1–11

Pieper GM, Langenstroer P, Siebeneich W (1997) Diabetic-induced endothelial dysfunction in rat aorta: role of hydroxyl radicals. Cardiovasc Res 34:145–156

Sand C, Michel MC (2014) Bradykinin contracts rat urinary blader largely independent of phospholipase C. J Pharmacol Exp Ther 348:25–31

Schneider T, Fetscher C, Krege S, Michel MC (2004a) Signal transduction underlying carbachol-induced contraction of human urinary bladder. J Pharmacol Exp Ther 309:1148–1153

Schneider T, Hein P, Michel MC (2004b) Signal transduction underlying carbachol-induced contraction of rat urinary bladder. I. Phospholipases and Ca^{2+} sources. J Pharmacol Exp Ther 308:47–53

Schneider T, Hein P, Bai J, Michel MC (2005a) A role for muscarinic receptors or rho-kinase in hypertension associated rat bladder dysfunction? J Urol 173:2178–2181

Schneider T, Hein P, Michel-Reher M, Michel MC (2005b) Effects of ageing on muscarinic receptor subtypes and function in rat urinary bladder. Naunyn Schmiedeberg's Arch Pharmacol 372:71–78

Schneider T, Fetscher C, Michel MC (2011) Human urinary bladder strip relaxation by the ß-adrenoceptor agonist isoprenaline: methodological considerations and effects of gender and age. Front Pharmacol 2:11

Stevens LA, Sellers DJ, McKay NG, Chapple CR, Chess-Williams R (2006) Muscarinic receptor function, density and G-protein coupling in the overactive diabetic rat bladder. Auton Autacoid Pharmacol 26:303–309

Török J, Zemancikova A, Kocianova Z (2016) Interaction of perivascular adipose tissue and sympathetic nerves in arteries from normotensive and hypertensive rats. Phys Rep 65(Suppl. 3):S391–S399

Weber LP, Chow WL, Abebe W, MacLeod KM (1996) Enhanced contractile responses of arteries from streptozotocin diabetic rats to sodium fluoride. Br J Pharmacol 118:115–122

Wyse DG (1980) On the "normalization" of active developed force of isolated helical strips of muscular and elastic arteries for variation in wall thickness. J Pharmacol Methods 4:313–326

Yang Q, Fujii W, Kaji N, Kakuta S, Kada K, Kuwahara M, Tsubone H, Ozaki H, Hori M (2018) The essential role of phospho-T38 CPI-17 in the maintenance of physiological blood pressure using genetically modified mice. FASEB J 32:2095–2109

Yesilyurt ZE, Erdogan BR, Karaomerlioglu I, Müderrisoglu AE, Michel MC, Arioglu Inan E (2019) Urinary bladder weight and function in a rat model of mild hyperglycemia and its treatment with dapagliflozin. Front Pharmacol 10:911