Northwestern Medicine Feinberg School of Medicine | MSTP | BPC | Microsurgery and Preclinical Research Core | Northwestern University Feinberg School of Medicine Comprehensive Transplant Center | Wertheim Lab Organ Regeneration and Tissue Engineering

Multivariate description of gait changes in a mouse model of peripheral nerve injury and trauma

Bilal A. Naved[1, 2, 3], Shuling Han[1, 2], Kyle M. Koss[1, 2], Mary J. Kando[4], Jiao-Jing Wang[1,2], Craig Weiss[4], Maya G. Passman[5], Jason A. Wertheim[1,2] , Yuan Luo, Zheng J. Zhang[1,2]

[1]Dept. of Surgery, [2]Comprehensive Transplant Center, Feinberg School of Medicine, [3]Dept. of Biomedical Engineering, [4]Behavioral Phenotyping Core, [5]Barnard College, [6]Division of Health and Biomedical Informatics, Dept. of Preventive Medicine.

Background

- ~2 million people living with limb loss in the United States and >20 million with peripheral nerve injury of some form (Ziegler-Graham 2008).
- Experimental treatments like limb transplantation (VCA) or tissue engineering materials must be evaluated in animals.
- Methods of evaluating gait and neuromotor function in animals are univariate.
- However, treadmill gait monitoring systems generate multidimensional data.
- We found no rigorous multivariate evaluation of rodent treadmill gait in the literature.
- This absence spans across all models (e.g. Diabetic Neuropathy, ALS, MS, etc.)

Hypotheses

- Multivariate gait analysis will reveal biologically consistent relationships.
- There are latent factors that help intuitively understand gait.
- Not all measurable features are relevant for characterizing gait, there is a subset
- Using this subset to train models will be more accurate than using all features
- These models will be able to distinguish between different gait phenotypes, like a human eye can just by watching
- This can all be done using a limb transplant model developed by the CTC's microsurgery core (**Fig. 1**) in conjunction with the DigiGait in the BPC (**Fig. 2**).

Experimental Design

- We modeled increasing neuromusculoskeletal damage.
- Groups = control, nerve transection, limb tx.
- The microsurgery core has published a limb tx model that was used (**Fig. 1**).
- 14 animals in each group.
- After a 2-week recovery period animals were video taped walking on DigiGait.

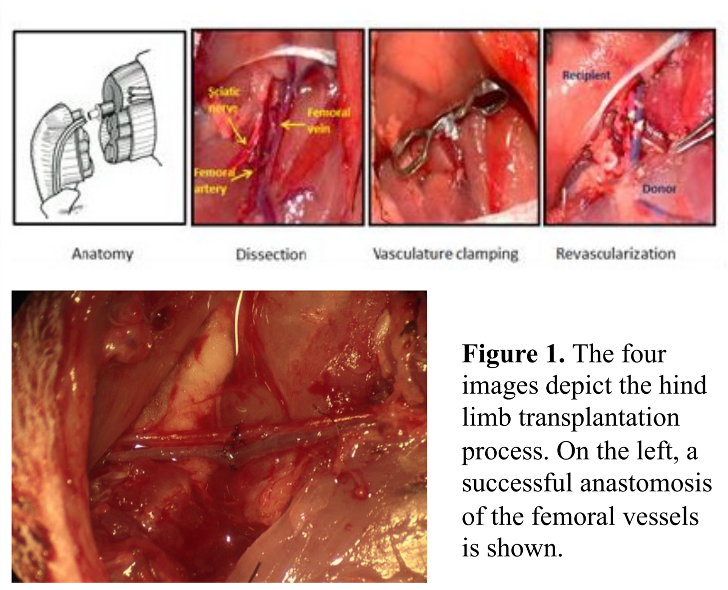


Figure 1. The four images depict the hind limb transplantation process. On the left, a successful anastomosis of the femoral vessels is shown.

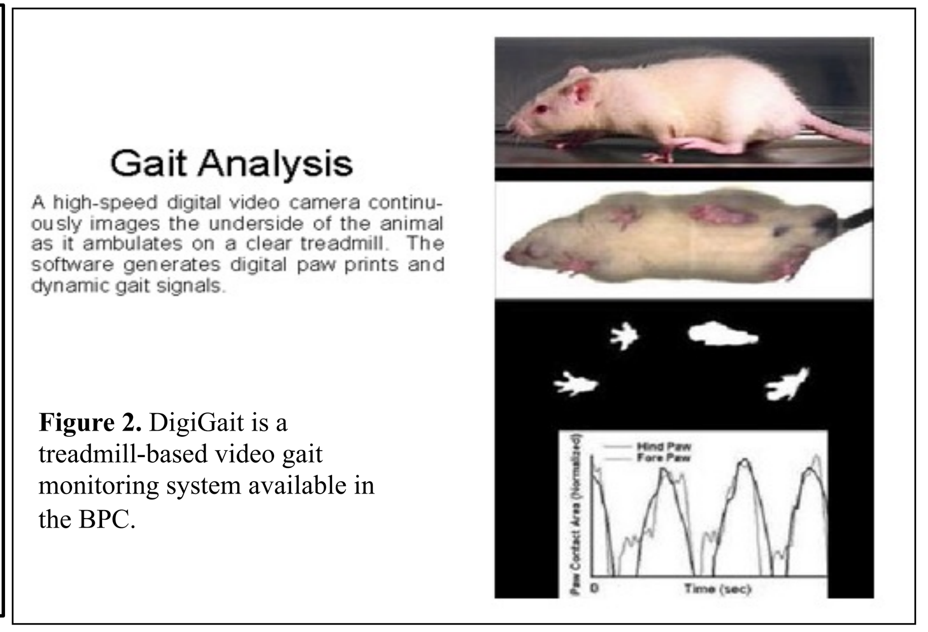
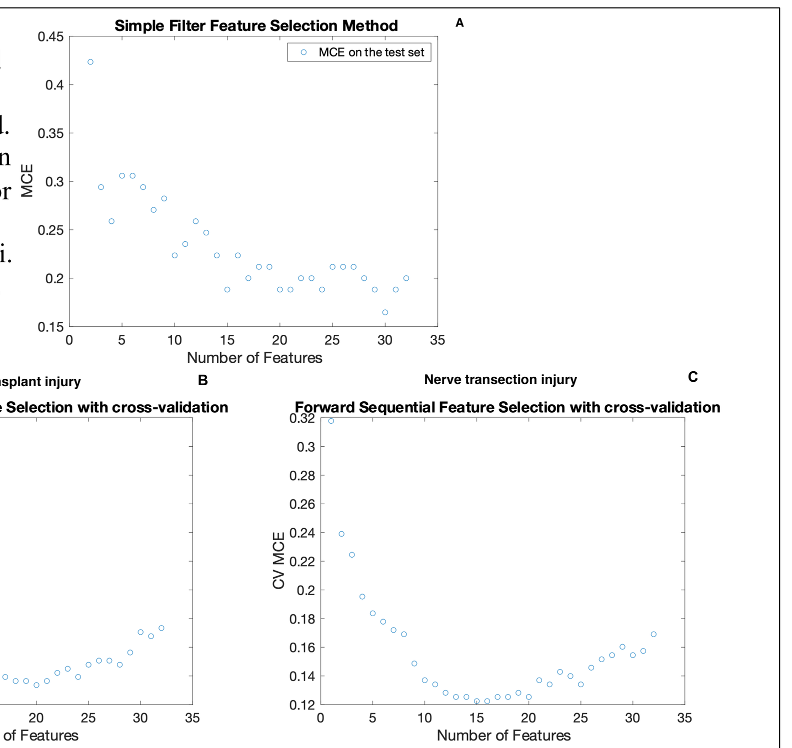


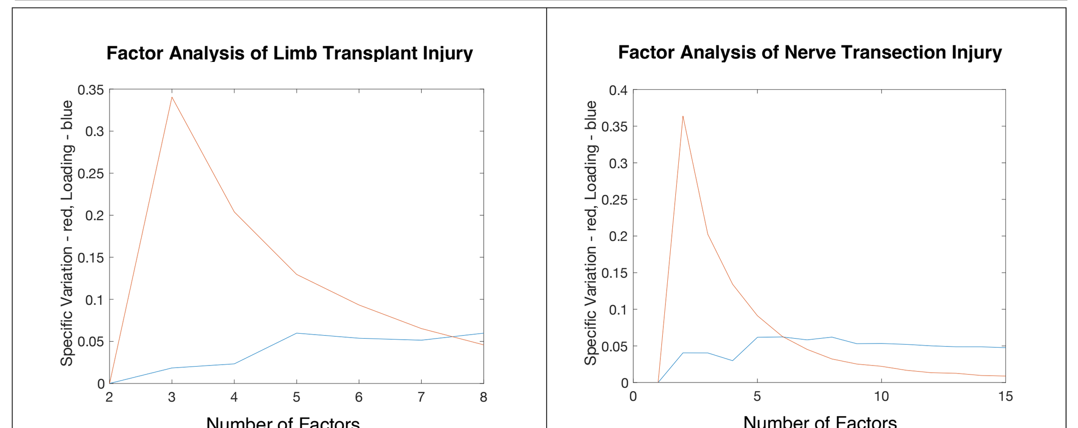
Figure 2. DigiGait is a treadmill-based video gait monitoring system available in the BPC.

Multivariate Characterization

- Univariate and multivariate dimensionality reduction were compared.
- This included multiple hypothesis testing and forward selection resp.
- The performance of a training algorithm was used to compare the two.
- Multivariate feature selection led to 8% greater misclassification error.
- Moreover, univariate selection appeared to yield overfitting (**Fig. 3**).
- Factor Analysis was conducted to better understand relationships.
- Six factors seemed to max loading and minimize variance (**Fig. 4**).
- Upon examining the components of each of the six factors, biologically consistent relationships were observed (**Table 1**).
- E.g: alterations to the anatomical configuration of the limb in tx **and**
- Abberant fine motor function due to peripheral nerve injury.

Figure 3. Univariate feature selection as well as multivariate feature selection was conducted. Notice the erratic pattern of misclassification error via univariate vs. the smooth pattern via multi. Multi also demonstrates how more features does not = more accuracy





Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
0.6948	0.0520	0.0605	0.0493	**0.8459**	-0.1463
-0.0093	-0.0031	0.0508	-0.0042	**0.9544**	-0.1260
0.5054	-0.0251	0.0474	-0.0412	-0.4782	**-0.7868**
0.6381	0.0397	-0.0617	0.1456	-0.1534	**0.5137**
0.0685	0.0379	-0.0865	0.1026	-0.2589	**0.8892**
0.2070	0.0715	**1.0258**	0.0853	0.0982	-0.0917
-1.0131	-0.0519	0.0106	-0.0794	-0.0896	0.0529
-0.0207	-0.0110	-0.0468	0.1400	0.0616	-0.3058
0.0925	0.0201	-0.0057	**0.7097**	0.1510	-0.0117
-0.2135	0.1322	-0.1985	0.3518	0.1155	-0.2952
0.1155	**1.0858**	0.3534	-0.1398	0.0327	0.0245
-0.0096	**0.8788**	-0.2840	-0.1811	0.0031	0.0471
0.0573	0.4291	-0.1908	0.0458	-0.1100	0.0157
0.1787	-0.1100	0.1257	0.5004	-0.1206	0.3665
-0.1044	0.2059	-0.1576	0.2591	0.0011	0.0418
0.0401	0.0661	-0.1442	0.0564	-0.2297	-0.0957

Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
0.5563	-0.0938	0.0153	0.0505	**0.9734**	-0.0441
-0.2529	-0.0996	-0.0317	0.0444	**0.8424**	-0.0004
0.5662	**-0.9102**	-0.0685	0.0175	-0.2256	0.1141
0.6907	0.5912	0.0825	0.0341	-0.0810	-0.0916
0.1294	**0.9102**	0.0643	-0.0328	-0.3042	-0.1259
0.1225	-0.0071	0.0326	**1.0162**	0.1015	0.0121
-1.0465	0.0337	-0.0129	-0.0162	-0.2239	0.0407
0.0257	-0.2207	-0.0117	0.0206	-0.0541	-0.0224
0.1443	0.2442	-0.0670	-0.0703	-0.0184	0.1192
-0.0006	-0.0925	0.1947	-0.1498	0.0729	0.0887
0.0717	0.0460	**1.0585**	0.3507	0.0004	0.0151
-0.0262	0.0183	**0.8864**	-0.2369	-0.0207	-0.0072
0.0159	-0.0168	0.4160	-0.2858	0.0698	0.1243
0.1088	0.4399	-0.1083	0.0483	-0.0222	**0.6586**
-0.1325	-0.0458	0.1472	-0.0925	-0.0045	**0.6317**
0.1341	-0.0967	0.0480	-0.1653	-0.1760	0.0432

Figure 4. Factor Analysis allows one to identify hidden (latent) factors in this case 6ish

Table 1. Factor analysis results of limb transplant data in contrast to only nerve transection injury data showing the features that make up each factor

Factor Analysis Results of Data From Pathological Gait Due To Limb Transplantation					
Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
Swing	SL Var	StrideLength	**Absolute Paw Angle**	Swing	Brake
Propel	Stride Length CV	Belt Speed	**Midline Distance**	%SwingStride	Propel
Stride Frequency					%PropelStance
Factor Analysis Results of Data From Pathological Gait Due To Nerve Transection					
Factor 1	**Factor 2**	**Factor 3**	**Factor 4**	**Factor 5**	**Factor 6**
Propel	Brake	SL Var	Stride Length	Swing	**Paw Area at Peak Stance in sq. cm**
Stride Frequency	%Propel Stance	Stride Length	Belt Speed	%SwingStride	**Paw Area Variability at Peak Stance in sq. cm**
					Min dA/dT
Latent factors that describe gait in the setting of traumatic nerve injury					
Relationship of stride length and speed	**Stride length and its variation**		**Measures of the paw**	**Measures of symmetry**	**Phases of the stride**

Notice key similarities (unbolded) and differences (**bolded**) between the two datasets. The bottom two rows describe the latent or implied factors that make sense of the various groups of observable measures shown above. These latent factors are an interpretation of the factor groupings shown above. The factor groupings were identified via factor analysis, which quantifies the relationship between individual, measurable features and can be used to identify the most highly grouped features.

Table 2. Evaluating the performance of 4 different model architectures in distinguishing between healthy and pathological phenotypes of gait

A. Performance in Distinguishing Peripheral Gait Deficit from Healthy Gait	**Accuracy**	**Precision**	**Recall**	**F-Score**
Random Forest	0.7294	0.7560	0.7521	0.7492
Discriminant Analysis	0.7477	0.8022	0.7634	0.7742
Support Vector Machine	0.7744	0.8108	0.7915	0.7948
Regression	0.7868	0.8570	0.7826	0.8130
Ensemble	0.9099	0.9283	0.9086	0.9165
B. Performance in Distinguishing Between Two Phenotypes of Peripheral Gait Deficit: Limb Transplant from only Nerve Transection	Accuracy	Precision	Recall	F-Score
Random Forest	0.6435	0.7072	0.6852	0.6878
Discriminant Analysis	0.7165	0.7827	0.7188	0.7388
Support Vector Machine	0.6987	0.7806	0.7341	0.7457
Regression	0.7237	0.7882	0.7270	0.7456
Ensemble	0.8780	0.9263	0.8781	0.8984

Using the identified features from feature selection + factor analysis 4 different model architectures were evaluated for their accuracy, precision, recall, and F-score in their ability to distinguish (**A**) healthy gait from gait deficit due to peripheral nerve injury and (**B**) gait deficit due limb transplantation from gait deficit due to total nerve transection

Machine Learning Classification of Gait States

- We used an 80-20 training-testing split with 10-fold cross validation.
- Five model architectures were evaluated (**Table 2**).
- Scores from the highest performing model architecture were plotted to evaluate the separation of their spreads + means (**Fig. 5**).

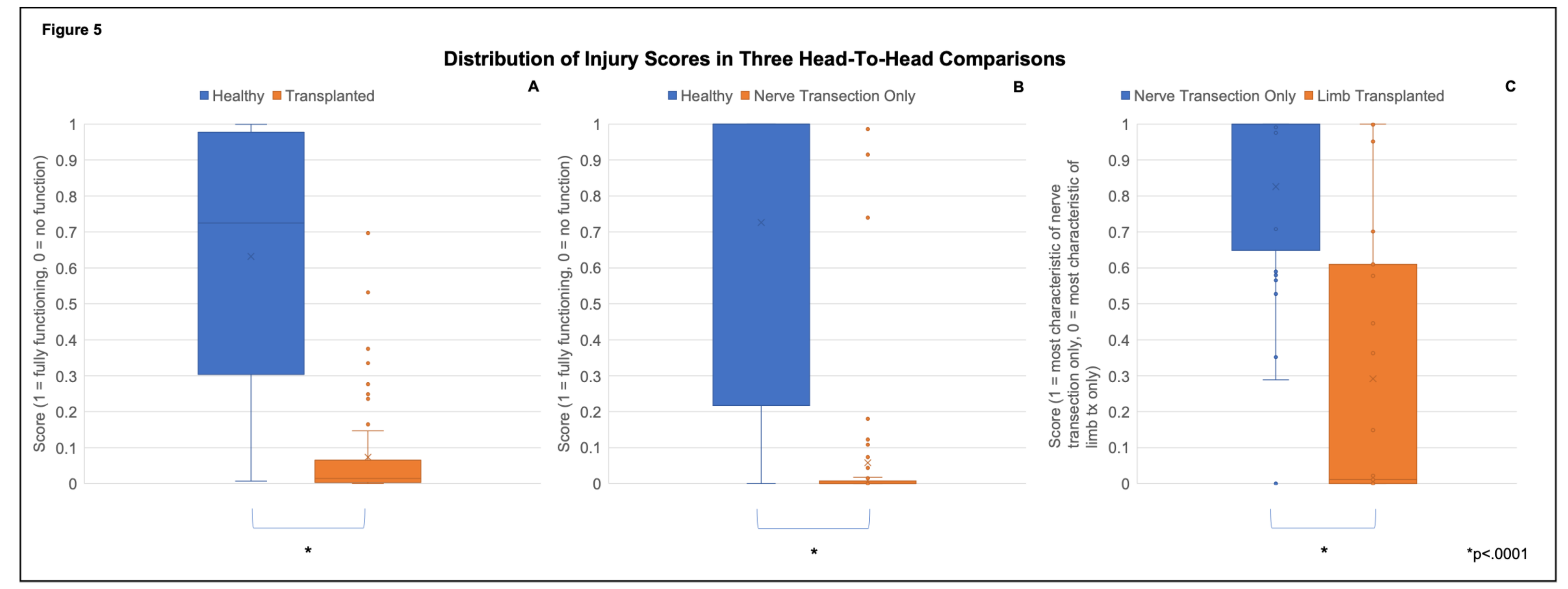


Figure 5. This visualizes the distribution of scores when the trained models were used to distinguish between different gait phenotypes. The models were able to distinguish between two distinct phenotypes with high statistical significance.

Results, Discussion, and Conclusion

- 16 features maximized predictive ability in contrast to the 32 available.
- This demonstrates the value of feature selection.
- Univariate selection performed 6% less accurately than multivariate selection.
- Multivariate selection did not help us understand exact relationships.
- Using factor analysis we identified the 6 latent factors most responsible for describing gait in the context of peripheral nerve injury and trauma.
- These 6 factors were composed of the 16 features and were biologically consistent revealing insights and hypotheses that were unable to without.
- Using the identified features, various models were trained.
- Ensemble-based classifiers achieved >90% classification accuracy with similarly high precision, recall, and F-score.
- Moreover, these classifiers were able to distinguish between the two different etiologies of gait with almost 90% accuracy as well.
- This is the first example of multivariate description and rodent gait classification between two different etiologies of deficit.
- This same technique could be used to make direct comparisons between completely different gait states or interventions.