**Benchmark datasets for bilateral lower-limb neuromechanical signals from wearable sensors during unassisted locomotion in able-bodied individuals**

Blair Hu\*1,2, Elliott Rouse3, Levi Hargrove1,2,4

1Center for Bionic Medicine, Shirley Ryan AbilityLab, Chicago, IL, USA

2Department of Biomedical Engineering, Northwestern University, Evanston, IL, USA

3Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI, USA

4Department of Physical Medicine and Rehabilitation, Northwestern University, Chicago, IL, USA

**\* Correspondence:**Blair Hu  
blairhu@u.northwestern.edu

Keywords: gait, locomotion, biomechanics, electromyography, benchmark

# Introduction

The field of assistive robotics has experienced rapid growth in the number of and capabilities of wearable lower limb assistive devices, which include robotic exoskeletons, orthoses, and prostheses. These devices have shown promising potential to restore motor function to individuals with gait impairments by providing locomotion assistance. Although many devices have already demonstrated impressive performance in a variety of real world conditions, comparing their performance objectively and improving their controllability remain challenging for several reasons. First, the outcome measures (*e.g.* joint kinematics, metabolic cost, clinical scores, prediction accuracy) used by studies demonstrating improved walking ability with an assistive device are not consistent. Second, many studies only use treadmill walking or do not collect data from a variety of locomotor activities due to constraints in a device’s mechatronic design and/or control system. Third, many devices are in the process of commercialization so testing data are seldom shared with the research community. In addition, many devices implement their own unique control frameworks which are not generic enough to conveniently implement on other hardware. Therefore, we expect improving access to device-agnostic neuromechanical signals during walking-related activities (from which researchers could develop and test novel control strategies before implementation on hardware) will be valuable to the field of wearable lower limb assistive devices.

Meanwhile, many benchmarks for the biomechanics of able-bodied human locomotion without an assistive device have already been established, some of which are publicly available. The gold standard for high-resolution biomechanical gait analysis is marker-based optical motion capture with ground reaction force measurement. Decades ago, seminal work from Winter (1983) used these techniques to introduce an inter-subject biomechanical analysis of level ground walking at different speeds. Their normative gait dataset includes electromyography (EMG) and joint kinematic and kinetic patterns and has since been expanded by other researchers to include more subjects and strides (*e.g.* Kadaba, Ramakrishnan, & Wootten, 1990; Kirtley, 2014). The steady-state biomechanics of other common locomotor activities such as ascending and descending stairs and sloped surfaces of different geometries have also been reported in separate studies using similar techniques but those data are not as accessible to researchers (*e.g.* Franz, Lyddon, & Kram, 2011; Lay, Hass, & Gregor, 2006; Lay, Hass, Richard Nichols, & Gregor, 2007; McFadyen & Winter, 1988; Riener, Rabuffetti, & Frigo, 2002; Protopapadaki, Drechsler, Cramp, Coutts, & Scott, 2007).

Human locomotion is most accurately quantified by joint kinematics, kinetics, and EMG using traditional laboratory-based instrumentation and techniques developed for biomechanical gait analysis. However, the exciting potential of wearable lower limb robotics lies in its promise to bring these devices closer to everyday life, where alternative techniques are required to more ubiquitously measure neuromechanical signals during walking-related activities. Methods to more freely measure human movement have been developed in the field of human activity recognition (HAR), which aims to use continuous streams of sensor data to recognize and monitor common activities of daily living such as sleeping, walking, exercising, and manipulating objects. As a result, HAR has produced an abundance of publicly available datasets. However, although these repositories are valuable because they contain many different types of activity information from many subjects, they are not very suitable for more systematic characterization of normal locomotion. Sometimes, HAR datasets are collected from impaired populations or during more natural but complex combined movements for which the ground truth activity is more ambiguous. Some datasets are collected using minimal instrumentation (*e.g.* smartphone only) which is convenient but incomplete. In contrast, others rely on non-portable instrumentation (*e.g.* optical motion-capture or video) which is highly accurate but not representative of biomechanical signals accessible for controlling a device in a more ecological setting. Also, many only contain single modalities (*e.g.* kinematics but no EMG) and/or use lower sampling rates that may be insufficient for certain online control schemes.

To the best of our knowledge, there still does not exist a publicly available database of kinematic and EMG data simultaneously recorded from wearable sensors as able-bodied individuals freely transition between several distinct locomotor activities. To address some of these aforementioned limitations and provide relevant reference data for researchers in the field of wearable lower limb assistive devices, we introduce a device-agnostic benchmark dataset of bilateral neuromechanical signals called ENABL3S(**EN**cyclopedia of **A**ble-bodied **B**ilateral **L**ower **L**imb **L**ocomotor **S**ignals). The dataset contains bilateral EMG and joint and limb kinematics recorded from wearable sensors for 10 able-bodied individuals as they freely transitioned between sitting, standing, and several walking-related activities (level ground, stair ascent/descent, and ramp ascent/descent). Although these data are not intended to replace existing benchmarks for biomechanical gait analysis, we believe they still fill a gap between those benchmarks and HAR datasets by providing richer neuromechanical data collected from wearable sensors using a unified protocol for several distinct locomotor activities. In this data report, we summarize our methods for instrumenting subjects, collecting data, and post-processing for artifact removal and gait segmentation. We also present a summary of the types of locomotor activities and transitions captured by our protocol, validate our results, and conclude with suggestions for how other researchers in the field may benefit from this dataset.

# Material and Methods

**2.1 Instrumentation setup**

Ten healthy, able-bodied subjects (7 male, 3 female; 25.5 ± 2 years; 174 ± 12 cm; 70 ± 14 kg) without any gait impairments were recruited and completed the following protocol between January and February 2017. Before walking, subjects were instrumented with wearable sensors to measure bilateral lower limb muscle activity and joint and limb kinematics. Electromyographic (EMG) signals were recorded using bipolar surface electrodes (DE2.1; Delsys, Boston, MA, USA) from the same seven muscles in each leg: tibialis anterior (TA), medial gastrocnemius (MG), soleus (SOL), vastus lateralis (VL), rectus femoris (RF), biceps femoris (BF), and semitendinosus (ST). These muscles were chosen because they are in part responsible for hip and knee flexion/extension and ankle plantarflexion/dorsiflexion, movements that are commonly assisted by wearable devices. They are also relatively easy to target when facing the subject from in front and behind. The muscle sites were prepared by removing excess hair and the skin was cleaned by mildly scrubbing with an alcohol wipe. Sensors were attached to the skin with a double-sided adhesive. Electrode placement was guided by palpation according to the Surface ElectroMyoGraphy for the Non-Invasive Assessment of Muscles (SENIAM) standards and verified by having subjects perform maximum voluntary contractions (MVC). Subjects performed three repetitions of ankle dorsiflexion/plantarflexion and knee flexion/extension for both legs. EMG signals were amplified by 1000x, hardware band-pass filtered between 20-450 Hz (Bagnoli 16, Delsys), and sampled at 1 kHz.

Joint kinematic signals (sagittal plane only) were recorded using electrogoniometers (SG150; Biometrics Ltd, Newport, UK) placed on the knee and ankle and sampled at 500 Hz. At the beginning of trials, the goniometers were zeroed while the subject was in the upright standing position. 6-DOF (tri-axial accelerometer and gyroscope) inertial measurement units (IMU) were placed bilaterally on the subjects’ thigh (below RF) and shank (adjacent to TA) and sampled at 500 Hz (MPU-9250; Invensense, San Jose, CA, USA). Goniometers and IMU’s were secured to the subject using a combination of double-sided adhesive, elastic straps, and Coban self-adherent wrap. Another IMU was placed in a custom manufactured holster (tilted 20° from vertical) and worn around the waist with a belt. All signals were simultaneously recorded with a custom 16-bit data acquisition device that permitted multi-rate sampling. To facilitate integration with our custom data acquisition software, all wearable sensors were used in a tethered setup; as a drawback, fully instrumenting each leg took up to an hour. The full instrumentation setup with IMU orientations is shown for a representative subject in Figure 1.

**2.2 Data collection protocol**

In an experimental session, each subject was barefoot and completed approximately 25 repetitions of a circuit consisting of sitting (S), standing (St), level ground walking (LW), ascending/descending a ramp with a 10° slope (RA/RD), and ascending/descending a four-step staircase (SA/SD) step-over-step. These activities were chosen because they encompass the different types of terrain likely encountered in community ambulation and were completed as a circuit in a 20 x 30 ft. room for practicality and for increasing the number of repetitions. A platform (30 in. tall) joined the staircase (7.75 in. rise, 10 in. run) and ramp (14 ft. long) to allow all possible transitions between these activities. Data from each circuit were divided into two segments and recorded as separate trials. Odd-numbered trials consisted of S 🡪 St 🡪 LW 🡪 SA 🡪 LW 🡪 RD 🡪 LW 🡪 St 🡪 S. Even-numbered trials consisted of S 🡪 St 🡪 LW 🡪 RA 🡪 LW 🡪 SD 🡪 LW 🡪 St 🡪 S. The total distance walked for each continuous segment was approximately 45 ft. Trials during which sensors needed to be repositioned or the tether became tangled were excluded. Subjects were instructed to freely transition between locomotor activities at their self-selected speed and breaks were routinely administered to avoid fatigue. The experimenter labeled the true locomotor intent of the subject using a key fob. Data collection took up to 2 hours.

**2.3 Post-processing**

Heel contact and toe off gait events for each leg were reliably identified by finding peaks in the mean-subtracted and low-pass filtered (1st order Butterworth, 6 Hz) sagittal plane angular velocity (GY) of the shank segment using a threshold-based method similar to (Maqbool et al., 2016). Briefly, the largest peaks in angular velocity were first used to identify mid-swing events. Toe off events were identified by searching for peaks before each mid-swing event. Heel contact events were identified by searching for peaks after each preceding mid-swing event. Event switches were initially placed beneath the heel and first metatarsal of each foot but they triggered many false negatives and positives in our setup perhaps due to mechanical wear and/or foot placement on the staircase. Therefore, they were only used for validating the IMU-based segmentation technique. Gait events corrupted by motion artifacts (*i.e.* pauses and trips) were excluded. EMG signals were high-pass filtered (6th order Butterworth) at 20 Hz, low-pass filtered (6th order Butterworth) at 350 Hz, and notch-filtered (6th order Butterworth, 6 Hz width) at 60, 180, 300 Hz to attenuate motion artifact and ambient interference. Goniometer and IMU signals were low-pass filtered (6th order Butterworth) at 10 and 25 Hz, respectively. Joint velocities were indirectly computed by taking the central-difference numerical derivative of the joint position, and added to the goniometer channels.

All signals were segmented into analysis windows beginning 300 ms before each identified heel contact or toe off gait event. Four additional 300 ms analysis windows near each identified gait event (delayed by 30, 60, 90, and 120 ms relative to each event) were used. For each window, we extracted features previously used in intent recognition for control of a powered knee-ankle prosthesis. Features for goniometer and IMU channels included the mean, standard deviation, maximum, minimum, initial, and final values (Varol, Sup, & Goldfarb, 2010) (6 features/channel). Features for EMG signals included the mean absolute value (MAV), waveform length, number of zero crossings and slope sign changes, and the coefficients of a sixth-order autoregressive model (Hargrove, Englehart, & Hudgins, 2008; Huang, Englehart, Hudgins, & Chan, 2005) (10 features/channel). There were a total of 23 sensors (14 EMG, 4 goniometer, 5 IMU), 52 channels (14 EMG, 8 goniometer, 30 IMU), and 368 features (140 EMG, 48 goniometer, 180 IMU).

# Results

## The data are saved in CSV format in subject-specific folders and are available to download from Figshare at <https://figshare.com/s/04b1c4d0192dbbb1efa8>. Within each subject-specific folder, there is a metadata file, which catalogues the filenames, summary statistics (mean, standard deviation, minimum, maximum) of each goniometer channel, and signal-to-noise ratios (ratio of maximum to baseline RMS voltage) of each EMG channel for each circuit. Subject-specific folders also include folders for the raw and processed data from individual circuits, a folder containing the processed EMG signals from all muscles during MVC trials, and a folder containing the features extracted from the five different 300 ms analysis windows (beginning 300, 270, 240, 210, and 180 ms before the gait events identified for each leg). Data from individual circuits also contain a columns specifying the label of the true locomotor activity, the indices of heel contact and toe off gait events, and four-digit triggers denoting the outgoing and incoming locomotor activities and gait phases. The first row of each file is a header specifying the column order. The post-processed data from all trials are included for completeness although some trials include disturbances (*e.g.* pauses, trips, missed transitions), which are noted in the metadata file. However, only gait events from disturbance-free segments of trials are reported and used for feature extraction. Feature data also contain columns specifying the corresponding leg-phase (1, right heel contact; 2, right toe off; 3, left heel contact; 4, left toe off) and the four-digit trigger. The first row of the feature data is a header specifying the column order. Ipsilateral refers to the side on which the gait event was detected (*e.g.* the right leg for right heel contact and right toe off events).

## The overall composition of ENABL3S is shown in Table 1. For each subject, there were 530 ± 46 heel contact events and 536 ± 45 toe off events for each leg (mean ± standard deviation) after excluding transitions to or from standing. Additional subject information and an explanation of nomenclature and numbering are also included on Figshare.

# Discussion

ENABL3S represents a benchmark of bilateral lower limb neuromechanical signals recorded from able-bodied individuals using wearable sensors during unassisted locomotion. The purpose of introducing this dataset is not to replace existing benchmarks for biomechanical gait analysis of steady-state locomotion but to provide a publicly available set of rich biomechanical data from wearable sensors, representing a compromise between traditional techniques and methods from HAR. ENABL3S includes data from several distinct walking-related activities (with transitions), which we expect to be helpful for understanding patterns in normal locomotion and developing novel control strategies for wearable lower limb assistive devices.

In order to assess the validity of this dataset, we chose to compare our recordings (averaged across legs and all subjects) to previously reported biomechanical measurements of level walking because those data are most accessible. Due to movement out of the sagittal plane and skin deformation/relative motion of the ankle goniometer, our measurements of ankle position were not considered biomechanically accurate signals. Nonetheless, those signals may still be useful for developing control strategies because many devices do not reproduce physiological motion and/or use embedded joint encoders to sense relative ankle position. However, our measurements of knee position were more accurate when compared to previously reported data recorded using optical motion capture (McClelland, Webster, Feller, & Menz, 2011; Winter, 1983). The root-mean-square (RMS) error between ENABL3S and Winter (1983) was 5.9 and 9.5 degrees for stance and swing phases, respectively. The R2 values were 0.73 and 0.94 for stance and swing phases, respectively. Our reported values for knee range of motion (ROM) during stance and swing phases (flexion at initial contact: 10.9 ± 5.6 degrees; stance ROM: 6.8 ± 5.2 to 26.9 ± 5.7 degrees, swing ROM: 3.4 ± 5.4 to 58.0 ± 6.5 degrees) were also comparable to reported values (McClelland et al., 2011). Errors in position can be attributed to a combination of differences in walking speed, minor misalignment of the sensor with the axis of rotation, and skin deformation/relative motion. The knee position could also be estimated (perhaps more accurately) by subtracting the orientations of the shank and thigh IMU sensors. The patterns of EMG activation for ankle plantarflexor/dorsiflexor and knee flexor/extensor muscles were also qualitatively similar to those previously reported for unassisted overground walking at self-selected speed (Sylos-Labini et al., 2014; Winter, 1983). Knee position and EMG from TA, MG, BF, and VL aggregated across legs for all steady-state level walking steps for all subjects can be found in a supplementary document on Figshare. By confirming the accuracy of our measured kinematic and EMG signals, we also validate the IMU-based method for gait segmentation.

Although these data are not as high resolution as optical motion capture, they strike a balance between resolution of signals, breadth of activities represented, feasibility for online control schemes, and contribution to existing publicly available datasets for human locomotion. These data can be used for developing novel control strategies such as intent recognition (*i.e.* predicting future states based on signals detected before movement completion) and more specifically investigating sensor fusion techniques and machine learning approaches for feature extraction and classification (*e.g.* deep learning). These data can also be interpreted as a simulation of able-bodied individuals walking with a completely massless and transparent (*i.e.* perfectly backdrivable) device and can be used to derive a device-agnostic upper bound on control strategies such as intent recognition. The raw data reported here may also be useful for comparing the performance of alternative control systems, assessing inter-subject variability, and comparing user-based biomechanical signals collected from an impaired population or a population walking with an assistive device (*e.g.* knee orthosis, ankle-foot prosthesis) to unassisted normal locomotion.

# Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# Author Contributions

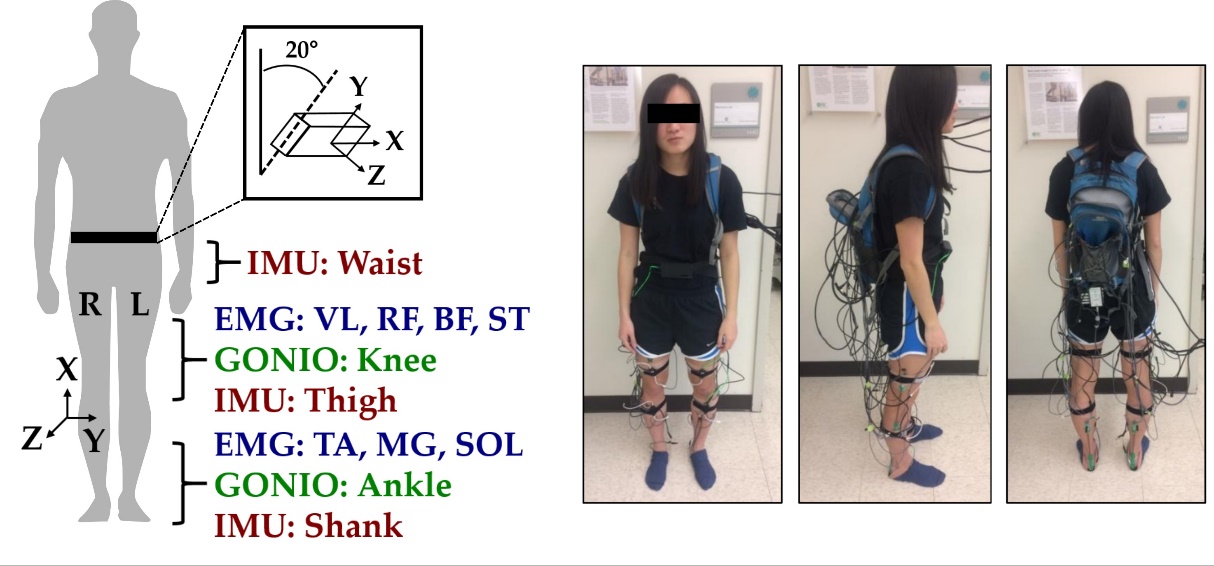
BH helped in conceiving the study concept, collecting, analyzing, and interpreting the data, and drafting the manuscript. LH helped in conceiving the study concept and interpreting the data, critically revising the manuscript for important intellectual content, obtaining funding, and supervising the study. ER helped in conceiving the study concept and interpreting the data, critically revising the manuscript for important intellectual content, obtaining funding, and supervising the study. All authors read and approved the final manuscript.

# Funding

The research was supported by USSOCOM Contract No. H92222-16-C-0111.

# Figures and Tables

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Table 1.** Characteristics of ENABL3S. The total number and proportion of gait events belonging to each type of locomotor activity are aggregated across all subjects. | | | | |
|  | **Transition to** | **Heel contact** | **Toe off** | **Total** |
| **Level walking**  **(LW)** | LW  RA  RD  SA  SD | 4523  240  240  239  248 | 4637  245  246  253  243 | 9160 (42.96%)  485 (2.27%)  486 (2.28%)  492 (2.31%)  491 (2.30%) |
| **Ramp ascent**  **(RA)** | RA  LW | 1408  243 | 1416  252 | 2824 (13.24%)  495 (2.32%) |
| **Ramp descent**  **(RD)** | RD  LW | 1757  239 | 1762  245 | 3519 (16.50%)  484 (2.27%) |
| **Stair ascent**  **(SA)** | SA  LW | 489  238 | 472  245 | 961 (4.51%)  483 (2.27%) |
| **Stair descent**  **(SD)** | SD  LW | 475  248 | 478  242 | 953 (4.47%)  490 (2.30%) |
|  |  | 10587 | 10736 | 21323 (100%) |



**Figure 1.** Instrumentation setup showing bilateral sensor placement. The orientation of the shank, thigh, and waist IMU’s are shown with coordinate axes. The subject provided written informed consent for the publication of this image.

# References

Franz, J. R., Lyddon, N. E., & Kram, R. (2011). Mechanical work performed by the individual legs during uphill and downhill walking. https://doi.org/10.1016/j.jbiomech.2011.10.034

Hargrove, L. J., Englehart, K. B., & Hudgins, B. (2008). A training strategy to reduce classification degradation due to electrode displacements in pattern recognition based myoelectric control. *Biomedical Signal Processing and Control*, *3*(2), 175–180. https://doi.org/10.1016/j.bspc.2007.11.005

Huang, Y., Englehart, K. B., Hudgins, B., & Chan, A. D. C. (2005). A Gaussian Mixture Model Based Classification Scheme for Myoelectric Control of Powered Upper Limb Prostheses. *IEEE Transactions on Biomedical Engineering*, *52*(11), 1801–1811. https://doi.org/10.1109/TBME.2005.856295

Kadaba, M. P., Ramakrishnan, H. K., & Wootten, M. E. (1990). Measurement of lower extremity kinematics during level walking. *Journal of Orthopaedic Research*, *8*(3), 383–392. https://doi.org/10.1002/jor.1100080310

Kirtley, C. (2014). CGA normative gait database. Retrieved from http://www.clinicalgaitanalysis.com/data/

Lay, A. N., Hass, C. J., & Gregor, R. J. (2006). The effects of sloped surfaces on locomotion: a kinematic and kinetic analysis. *Journal of Biomechanics*, *39*(9), 1621–8. https://doi.org/10.1016/j.jbiomech.2005.05.005

Lay, A. N., Hass, C. J., Richard Nichols, T., & Gregor, R. J. (2007). The effects of sloped surfaces on locomotion: an electromyographic analysis. *Journal of Biomechanics*, *40*(6), 1276–85. https://doi.org/10.1016/j.jbiomech.2006.05.023

Maqbool, H. F., Husman, M. A. B., Awad, M. I., Abouhossein, A., Mehryar, P., Iqbal, N., & Dehghani-Sanij, A. A. (2016). Real-time gait event detection for lower limb amputees using a single wearable sensor. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 5067–5070). IEEE. https://doi.org/10.1109/EMBC.2016.7591866

McClelland, J. A., Webster, K. E., Feller, J. A., & Menz, H. B. (2011). Knee kinematics during walking at different speeds in people who have undergone total knee replacement. *The Knee*, *18*(3), 151–155. https://doi.org/10.1016/j.knee.2010.04.005

McFadyen, B., & Winter, D. (1988). An Integrated Biomechanical Analysis of Normal Stair Asecent and Descent. *Journal of Biomechanics*, *21*(9), 733–744. https://doi.org/10.1016/0021-9290(88)90282-5

Protopapadaki, A., Drechsler, W. I., Cramp, M. C., Coutts, F. J., & Scott, O. M. (2007). Hip, knee, ankle kinematics and kinetics during stair ascent and descent in healthy young individuals. *Clinical Biomechanics*, *22*(2), 203–210. https://doi.org/10.1016/j.clinbiomech.2006.09.010

Riener, R., Rabuffetti, M., & Frigo, C. (2002). Stair ascent and descent at different inclinations. *Gait and Posture*, *15*(1), 32–44. https://doi.org/10.1016/S0966-6362(01)00162-X

Sylos-Labini, F., La Scaleia, V., d’Avella, A., Pisotta, I., Tamburella, F., Scivoletto, G., … Ivanenko, Y. P. (2014). EMG patterns during assisted walking in the exoskeleton. *Frontiers in Human Neuroscience*, *8*, 423. https://doi.org/10.3389/fnhum.2014.00423

Varol, H. A., Sup, F., & Goldfarb, M. (2010). Multiclass Real-Time Intent Recognition of a Powered Lower Limb Prosthesis. *IEEE Transactions on Biomedical Engineering*, *57*(3), 542–551. https://doi.org/10.1109/TBME.2009.2034734

Winter, D. A. (1983). Biomechanical motor patterns in normal walking. *Journal of Motor Behavior*, *15*(4), 302–30. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15151864