Digital Phenotyping in Patients with Spine Disease: A Novel Approach to Quantifying Mobility and Quality of Life.

David J. Cote, BS[1,*], **Ian Barnett, PhD**[2,*], **Jukka-Pekka Onnela, PhD**[2,†], **Timothy R. Smith, MD, PhD, MPH**[1,†]

[1]Computational Neuroscience Outcomes Center, Brigham and Women's Hospital, Harvard Medical School, Boston, MA;

[2]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA

Abstract

Objective: To identify trends in mobility and daily pain levels among a cohort of patients with clinically diagnosed spine disease.

Methods: Participants with spine disease were enrolled from a general neurosurgical clinic and installed a smartphone application (Beiwe™) designed for digital phenotyping to their personal smartphone. This application collected passive meta-data on a minute-to-minute basis, including GPS, WiFi, accelerometer, text and phone logs, and screen on/off time. The application also administered daily Visual Analogue Scale pain surveys. A Linear Mixed Model framework was used to test for associations between self-reported pain and mobility/sociability from passively collected data.

Results: 105 patients were enrolled with a median follow-up time of 94.5 days; 55 patients underwent a surgical intervention during follow up. Weekly pain survey response rate was 73.2%. By the end of follow up, the mean change in pain for all patients was −1.3 points (4.96 at the start of follow up to 3.66 by the end of follow up). Increased pain was significantly associated with reduced patient mobility as measured by three daily GPS summary statistics (average flight length, maximum diameter travelled, total distance travelled).

Conclusions: Patients with spine disease who report higher pain have reduced mobility as measured by passively collected smartphone GPS data. Smartphone-based digital phenotyping appears to be a promising and scalable approach to assess mobility and quality of life in patients with spine disease.

Corresponding Author: David J. Cote, BS, Brigham and Women's Hospital, Harvard Medical School, 60 Fenwood Road, Boston, MA 02115, Phone: 860 605 3401, david_cote@hms.harvard.edu.
*these authors contributed equally
†these authors contributed equally

Disclosures: The authors have no disclosures.

Keywords

spine; digital phenotyping; outcomes

Introduction

Few effective methods currently exist for objectively evaluating patient recovery after spine surgery. Among the most widely used and clinically evaluated are a broad set of measures known as Patient Reported Outcome Measures (PROMs), which often consist of validated questionnaires aimed at determining a patient's quality of life in response to surgical treatment [1]. These are paper-based or electronic surveys that require significant clinician and patient investment, necessitating multiple in-office follow-up appointments and regular patient contact. Typically, these are administered pre-operatively and at several points post-operatively, with exact administration varying by practice [1–3].

Unfortunately, questionnaire-based instruments have their limitations, both practical and conceptual [4,1]. First, they rely on patient recall at a particular moment in time, making them subject to bias. Second, they often require patients to return to clinic, or have a phone call or other interaction with a care provider. Third, they require the patient to make reporting assumptions regarding all events subsequent to the last encounter. If they are asked to report their pain, the patient must interpret if that means the average pain since the last visit, the average pain of the encounter day, or their pain at the current moment. Lastly, surveys provide a limited view of a patient's life, based solely on their ability to complete those tasks assessed in a particular questionnaire.

A recent study by Falavigna et al. highlighted some of these major limitations, demonstrating that among an international community of spine surgeons, 31.9% do not ever use PROMs for either research or clinical purposes [2]. Reasons for limited use included limited time in clinic for administration (57% of all respondents), difficult follow-up after discharge (36%), and lack of staff to assist in data collection (55%). In light of these limitations, novel methods for monitoring patient outcomes has been proposed, including digital phenotyping.

Digital phenotyping has been recently defined as the "moment-by-moment quantification of the individual-level human phenotype *in-situ* using data from personal digital devices," such as smartphones [5,6]. In this approach, subjects download and launch a smartphone application that collects both active data (such as surveys) and passive data (such as GPS data) from participants. This data is then used to study variations in patient behavior, including mobility (using GPS data), sociability (using text message and call logs), and sleep (using screen activity logs), among a variety of other measures. These objective measurements can be correlated with patient responses to phone surveys or any other data that might be available, including clinical examinations conducted at the clinic.

As smartphones have become ubiquitous, with ownership exceeding 77% of adults in the U.S.,[7] digital phenotyping for the purpose of improving patient outcomes through voluntary monitoring has become feasible and economical on large scales. Already, digital

phenotypes have been successfully used to link patient mobility to mood and depressive symptoms in a variety of clinical contexts [8–10]. While digital phenotyping is being used in the study and treatment of psychiatric disorders, its use for monitoring and aiding patients with debilitating physical conditions such as spine disease has not been attempted.

In this study, we report the first ever use of digital phenotyping in patients with spine disease in order to identify associations between behavioral data, passively collected from patients' smartphones, and daily self-reported pain. We report strong associations between self-reported pain and multiple aspects of patient mobility as measured by digital phenotyping.

Materials and Methods

Patient Recruitment

All patients included in this study were neurosurgical candidates with clinically diagnosed spine disease, seen in a general neurosurgical clinic. Enrollment began in June 2016 and continued through May 2017. Over the study period, 52.4% of patients underwent neurosurgical intervention (Table 1). Only adults were included in this study (18 years old), and patients who did not own a smartphone were excluded. The project was approved by the Institutional Review Board of Brigham and Women's Hospital, and patients provided expressed written consent to participate in this study. Upon entering the study, the research assistant helped each patient download the application onto their smartphone in clinic, and established each participant's user account. Patients were then given uniform instructions on how to operate the application, including keeping it running as a background application and responding to daily and weekly questionnaires.

Data Collection

Patients enrolled in the study installed the Beiwe application, which is part of the Beiwe research platform developed by a subset of the authors, onto their personal smartphone. The Android and iOS Beiwe applications collect both active and passive data from patients, and is accessible only with a unique username and password known only to the user.

Data collected by the current version of the Beiwe app includes global positioning satellite (GPS) data, accelerometer records, Bluetooth and Wi-Fi data, phone and call logs (Android only), and screen on/off time. These data streams are collected by the application at different customizable rates. In this study, we configured Beiwe to collect GPS data for 1 minute every 5 minutes and accelerometer data for 10 seconds every 10 seconds. It also collected anonymized phone call and text message logs, which included information about the timing of communication events, anonymized identifiers of communication partners, and message length for text message. The application did not record any of the actual content of phone calls or text messages. For modeling purposes, the collected raw sensor and phone usage data is represented in terms of daily summary statistics that capture salient features of each data stream (see Table 2). For each subject, the summary statistics are represented as a matrix, where the rows corresponding to different statistics and columns to different days. A visual representation of this data matrix for a single participant is shown in Figure 1.

In this study, patients were surveyed once per day at 5:00 p.m. EST with the following prompt: "Please rate your pain over the last 24 hours on a scale from 0 to 10, where 0 is no pain at all and 10 is the worst pain imaginable," with a sliding scale answer that ranged from 0–10. The Beiwe application encrypted the data as it was collected, stored it temporarily on the user's smartphone, and then periodically uploaded the data to a secure server via Wi-Fi [11].

Statistical Methods

Overall trends in pain over follow-up were evaluated as the difference between the first and last survey pain scores submitted over the course of follow-up.

In order to measure mobility, the raw GPS data for each patient was converted into a sequence of flights (straight-line movement) and pauses,[12] missing portions of data were imputed, and a variety of daily mobility summaries were produced [13]. Daily measures of sociability were summarized from call and text logs. These passively collected GPS mobility measures and call/text log sociability measures are listed in Table 2. With the aim of identifying how a patient's pain is related to mobility and sociability over the course of the same day and to avoid biased/inaccurate responses, we excluded late survey responses submitted after midnight (more than 7 hours after prompting).

Given the longitudinal nature of the data, to test for associations between daily self-report of pain and daily mobility and sociability, we used a Linear Mixed Model (LMM) [14]. For the daily mobility metrics, we ignored days that either had no GPS data or had no response to the pain survey. Similar to the analysis of Wang et al.,[8] we fit the following model for each of the 30 mobility and sociability passive data summaries: $y_{ij} = \beta_0 + \beta_1 X_{ij} + b_{i0} + b_{i1} X_{ij} + \epsilon_{ij}$, where the outcome y_{ij} is self-reported pain score on a 0–10 scale, X_{ij} is one of the 30 passive data summaries, and ϵ_{ij} is the normally distributed residual for the ith subject on their jth day of data collection. The fixed coefficients are the intercept (β_0) and slope (β_1), with random patient-specific coefficients (b_{i0}) and slopes (b_{i1}). This mixed-effects model allows for patient-specific relationships (intercept and slope) between the predictor and pain. Two-sided inference was performed on β_1 using a likelihood ratio test, and this modeling and testing procedure was repeated for each of the 30 passive data summary/features taking turns as the predictor X_{ij}. With 30 different and correlated tests being performed, we used the Generalized Higher Criticism to correct for multiple testing and identify statistically significant associations between passive data features and self-report of pain [15].

To measure the tendency of patients to stay home during prolonged periods of pain or discomfort, we estimated a daily probability of moving less than 1km, a somewhat arbitrary threshold that was selected to capture days of little to no movement within a reasonable margin of error. A Gaussian kernel, centered on the day of interest, averaged the indicator variables of whether or not the patient moved more than 1km across all days of the patient's follow up. Kernel averaging gives higher weight to days closer to the day of interest. This process was repeated to calculate a probability for each day in follow up.

Results

Over the enrollment period, 216 patients were approached for enrollment. Of these, 90 (42%) were immediately excluded because of lack of smartphone ownership. Of the remaining 126 patients, 15 (12%) could not recall necessary phone passwords for enrollment, four (3%) did not have their phone accessible on the day of the visit, and two (1%) declined consent over data security concerns; the remaining 105 (83%) patients were enrolled.

Demographics of included patients and relevant data collection metrics are shown in Table 1. On average, patients completed 43% of their daily surveys over the course of follow up; weekly pain survey response rate was 73.2%. Of the surveys that were completed, 71% were completed on the day they were administered (Figure 2). Median follow-up time was 3.15 months from enrollment, and 52.4% of patients underwent neurosurgical intervention during the study period.

Over the course of follow-up, the mean change in pain for all patients was −1.3 (from an average of 4.96 at the start of follow up to 3.66 by the end of follow up).

After correcting for multiple testing, average flight length, maximum diameter travelled, and total distance travelled were each statistically significantly associated with patient-reported pain. With the mobility trace of each patient broken into a sequence of flights (straight-line movement) and pauses (periods of stationarity), an average increase in self-reported pain by 0.1 was associated with 2.8-fold decrease in the average length of a patient's flights that same day (p=0.002),[16] a 5.1-fold decrease in a patient's maximum diameter, the largest distance between any two points the patient had been over the course of a day (p=0.004), and a 6.0-fold decrease in a patient's distance travelled over the course of a day (p=0.004). Associations between self-reported pain and various measures of mobility and sociability are shown in Table 3.

These cohort-level relationships between mobility and pain were averaged across all patients, and corresponded to β_1 from the above model [12]. The patient-specific relationships were captured by b_{i1} for the ith patient. Examples of patient-specific relationships between mobility and self-report of pain are visualized in Figure 3. Representative results for four patients using kernel averaging are shown in Figure 4.

Discussion

The limitations of clinical outcomes studies for patients with spine disease are well known and have been previously well described [17]. Comprehensive, quantitative evaluation of outcomes is impeded by recall bias and poorly-defined outcome measures. In clinical practice, patients with spine disease who are neurosurgical candidates are, at best, simply assessed with a validated PROMs instrument that can be used to track specific symptoms over time. In most cases, they are asked about their symptoms, their quality of life, and their ability to complete their activities of daily living. Over time, changes in overall symptoms are tracked based solely on these patient reports, and are not recorded in a systematic fashion [1]. In clinical research settings, PROMs surveys are administered at set intervals pre- and

post-operatively, usually including follow-up through at least one year after surgery. In typical clinical practice, they may not be used at all [2]. An international survey of spine surgeons by Falavigna et al. demonstrated that almost one third of spine surgeons did not use PROMs at all, for either research or clinical care [2].

Nevertheless, PROMs have frequently been used to assess outcomes after spinal surgery since they began to emerge in the early 2000s [1,2,18–22]. The results of these studies have often demonstrated improvement in patient quality of life and symptoms post-operatively, although limitations in follow-up and the quality of survey instruments have limited the generalizability of these findings [23,19].

In neurosurgical practice, the results of the limitations associated with PROMs and other classic survey instruments are serious. Often, neurosurgeons counsel patients on the benefits of surgical intervention based largely on subjective reports of pain and other symptoms, and imaging findings that may or may not correlate with severity of disease presentation. Without high-quality data demonstrating reduced quality of life with spine disease in a particular patient, or high quality data demonstrating the effectiveness of surgical intervention, decisions in the care of patients with spine disease are often made without strong evidence [17,24,25].

In this study, we report the first-ever application of smartphone-based passive data collection for the objective measurement of patient mobility in a cohort of 105 patients with spine disease. Over an average follow up period of 94.5 days, we used the Beiwe smartphone application and research platform to collect both active and passive data from these patients, in an effort to identify trends in self-reported pain and objective measures of mobility and sociability. In doing so, we demonstrated statistically significant associations between patient self-reported pain and patient mobility, but no associations between pain and sociability.

The relationship between pain and patient mobility has historically been difficult to study [26–31]. On the one hand, pain is an inherently subjective experience, and is affected by a multitude of factors, including the patient's disease state, current pain medication regimen, daily activities, and mental health. It changes constantly, is impossible to measure objectively, and varies from patient to patient. While it is reasonable to imagine that patients in significant pain would have reduced mobility, it is just as possible that those patients who increase their mobility suddenly may provoke significant pain. This multidirectional relationship makes identifying trends in patient quality of life based on pain scales difficult. With the advent of wearable pedometers that track mobility, some research has recently been performed that uses these more objective patient measures to evaluate mobility [32–34]. Unlike using smartphone based software, these studies typically require participants to wear an additional device, increasing the likelihood of patient non-compliance and subsequent missing data.

Similarly, the association between pain and other patient outcomes and sociability, social networks, or social support systems has also been difficult. Typically, studying these associations has involved in-depth interviews or PROMs-like paper or electronic

questionnaires completed by patients to identify networks of social support [35–37]. Outcomes can be tracked as usual and compared statistically with objective measures of patient social networks, with most studies demonstrating that patients with stronger social networks have reduced mortality and improved outcomes [38].

The novel method of in situ human research employed in this study, known as digital phenotyping, has the potential to revolutionize the study of these types of quality of life measures. Digital phenotyping has previously been used to assess patient mobility with mood and depressive symptoms [5,6,11]. The introduction of smartphone data-driven approaches to patient outcomes in psychiatric research has provided an objective grounding for understanding of patient behavior. We believe that among patients with spine disease, digital phenotyping has the potential to revolutionize patient care, which is increasingly driven by patient quality of life and personal outcome measures such as functional status and pain.

In this study of patients with spine disease, we identified three statistically significant relationships between pain and mobility. Specifically, patients who reported increased pain showed reduced mobility in that they traveled shorter total distances across each day, traveled within a narrower diameter, and took shorter average trips. Although we did not identify associations between pain and patient sociability, future studies focusing on other illnesses, including brain tumors, will attempt to identify relationships between disease and objective sociability, as measured by phone and text responsiveness.

Using non-invasive means through data collected from a patient's own personal smartphone, digital phenotyping allows for the moment-by-moment analysis of patient behavior. Using historical trends, analysis of data obtained from the Beiwe application can allow for identification of deviations in expected trends, possibly even allowing for clinical intervention. In patients with spine disease, digital phenotyping allows clinicians to track symptoms, such as pain and mobility, objectively over time. This could play an important role in operative planning, and in assessing patient response to neurosurgical procedures. In the future, digital phenotyping may even be useful for identifying adverse events during the post-operative period, such as the development of a surgical site infection, or changes in patient mental status.

The advantages of digital phenotyping are numerous, and include the ubiquity of personal cell phones, the low cost of installing and using the application, the outstanding granularity of data collected, and the ability to assess patient's self-reported pain on a daily basis. Additionally, digital phenotyping is non-invasive, and does not require patients to carry or use an additional device. In this way, digital phenotyping truly allows for analysis of patient behavior and quality of life in situ. Electronic survey response is easy and quick for patients to do, and has previously been shown to result in higher response rates than other modes, such as postal or telephone interviews [39,40].

Limitations of this study include the exclusion of patients without smartphones, which may result in underrepresentation of patients of lower socioeconomic status and the elderly, though recent data regarding cell-phone ownership demonstrate increased ownership across

the population [7]. Additionally, given that the collection of high-frequency sensor data causes some battery drainage, sensors need to be sampled according to a sampling scheme that unavoidably introduces some missingness by design. Some further missingness occurs because of human behavioral factors, such as individuals deactivating smartphone GPS. The missing data must then be imputed or accounted for in statistical calculations.

Despite these limitations, the current study nevertheless includes millions of data points on over 100 patients with spine disease, providing the first objective measurement of pain and mobility in this patient population based on *in situ* data rather than survey responses. Patients with spine disease who report increased pain demonstrate reduced mobility, as measured by passively collected GPS data from personal smartphones correlated with daily self-reported pain. The method of digital phenotyping is a novel way of objectively assessing quality of life in patients with spine disease on a moment-to-moment basis. Digital phenotyping has the potential to revolutionize the surgical care of patients with spine disease by providing objective measures of patient symptoms and functional status both before and after intervention.

Conclusions

Patients with spine disease who report higher pain have reduced mobility as measured by passively collected smartphone GPS data. Smartphone-based digital phenotyping appears to be a promising and scalable approach to assess mobility and quality of life in patients with spine disease.

Funding:

Abbreviations

|	
EST	Eastern standard time
GPS	global positioning satellite
LMM	linear mixed model
PROMS	Patient Reported Outcome Measures

References

1. Faraj SSA, van Hooff ML, Holewijn RM, Polly DW Jr., Haanstra TM, de Kleuver M: Measuring outcomes in adult spinal deformity surgery: a systematic review to identify current strengths, weaknesses and gaps in patient-reported outcome measures. European spine journal: official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society (2017). doi:10.1007/s00586-017-5125-4

2. Falavigna A, Dozza DC, Teles AR, Wong CC, Barbagallo G, Brodke D, Al-Mutair A, Ghogawala Z, Riew KD: Current status of worldwide use of patient-reported outcome measures (PROMs) in spine care. World neurosurgery (2017). doi:10.1016/j.wneu.2017.09.002

3. Nayak NR, Coats JM, Abdullah KG, Stein SC, Malhotra NR: Tracking patient-reported outcomes in spinal disorders. Surgical neurology international 6(Suppl 19), S490–499 (2015). doi: 10.4103/2152-7806.166892 [PubMed: 26605111]

4. Paulsen RT, Bouknaitir JB, Fruensgaard S, Carreon L, Andersen M: Prognostic Factors for Satisfaction After Decompression Surgery for Lumbar Spinal Stenosis. Neurosurgery (2017). doi: 10.1093/neuros/nyx298

5. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, Onnela JP: Utilizing a Personal Smartphone Custom App to Assess the Patient Health Questionnaire-9 (PHQ-9) Depressive Symptoms in Patients With Major Depressive Disorder. JMIR mental health 2(1), e8 (2015). doi: 10.2196/mental.3889 [PubMed: 26543914]

6. Torous J, Staples P, Onnela JP: Realizing the potential of mobile mental health: new methods for new data in psychiatry. Current psychiatry reports 17(8), 602 (2015). doi:10.1007/s11920-015-0602-0 [PubMed: 26073363]

7. Smith A: US smartphone use in 2015. Pew Research Center 1 (2015).

8. Wang R, Aung MS, Abdullah S, Brian R, Campbell AT, Choudhury T, Hauser M, Kane J, Merrill M, Scherer EA, Tseng VW: CrossCheck: Toward passive sensing and detection of mental health changes in people with schizophrenia. Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing, 886–897 (2016).

9. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, Mohr DC: Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. Journal of medical Internet research 17(7) (2015).

10. Faherty LJ, Hantsoo L, Appleby D, Sammel MD, Bennett IM, Wiebe DJ: Movement patterns in women at risk for perinatal depression: use of a mood-monitoring mobile application in pregnancy. Journal of the American Medical Informatics Association (2017).

11. Torous J, Kiang MV, Lorme J, Onnela J-P: New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR mental health 3(2) (2016).

12. Rhee I, Shin M, Hong S, Lee K, Kim SJ, Chong S: On the levy-walk nature of human mobility. IEEE/ACM transcations on networking (TON) 19, 630–643 (2011).

13. Khanolkar AR, Ljung R, Talback M, Brooke HL, Carlsson S, Mathiesen T, Feychting M: Socioeconomic position and the risk of brain tumour: a Swedish national population-based cohort study. Journal of epidemiology and community health (2016). doi:10.1136/jech-2015-207002

14. Mclean RA, Sanders WL, Stroup WW: A unified approach to mixed linear models. The American Statistician, 54–64 (1991).

15. Deb S, Pendharkar AV, Schoen MK, Altekruse S, Ratliff J, Desai A: The effect of socioeconomic status on gross total resection, radiation therapy and overall survival in patients with gliomas. Journal of neuro-oncology 132(3), 447–453 (2017). doi:10.1007/s11060-017-2391-2 [PubMed: 28258423]

16. Benjamini Y, Yekutieli D: The control of false discovery rate in multiple testing under dependency. Annals of Statistics, 1165–1188 (2001).

17. Forsth P, Olafsson G, Carlsson T, Frost A, Borgstrom F, Fritzell P, Ohagen P, Michaelsson K, Sanden B: A Randomized, Controlled Trial of Fusion Surgery for Lumbar Spinal Stenosis. The New England journal of medicine 374(15), 1413–1423 (2016). doi:10.1056/NEJMoa1513721 [PubMed: 27074066]

18. Devlin NJ, Parkin D, Browne J: Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. Health economics 19(8), 886–905 (2010). doi:10.1002/hec.1608 [PubMed: 20623685]

19. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY: Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. The spine journal: official journal of the North American Spine Society 8(6), 968–974 (2008). doi: 10.1016/j.spinee.2007.11.006 [PubMed: 18201937]

20. Brazier J, Jones N, Kind P: Testing the validity of the Euroqol and comparing it with the SF-36 health survey questionnaire. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation 2(3), 169–180 (1993).

21. BenDebba M, Heller J, Ducker TB, Eisinger JM: Cervical spine outcomes questionnaire: its development and psychometric properties. Spine 27(19), 2116–2123; discussion 2124 (2002). doi: 10.1097/01.brs.0000025729.35559.28 [PubMed: 12394923]

22. Auffinger BM, Lall RR, Dahdaleh NS, Wong AP, Lam SK, Koski T, Fessler RG, Smith ZA: Measuring surgical outcomes in cervical spondylotic myelopathy patients undergoing anterior cervical discectomy and fusion: assessment of minimum clinically important difference. PloS one 8(6), e67408 (2013). doi:10.1371/journal.pone.0067408 [PubMed: 23826290]

23. Baron R, Elashaal A, Germon T, Hobart J: Measuring outcomes in cervical spine surgery: think twice before using the SF-36. Spine 31(22), 2575–2584 (2006). doi:10.1097/01.brs. 0000240694.83621.52 [PubMed: 17047547]

24. Moojen WA, Van der Gaag NA: Minimally invasive surgery for lumbar spinal stenosis. European journal of orthopaedic surgery & traumatology: orthopedie traumatologie 26(7), 681–684 (2016). doi:10.1007/s00590-016-1828-1 [PubMed: 27659170]

25. Moojen WA, Arts MP, Jacobs WC, van Zwet EW, van den Akker-van Marle ME, Koes BW, Vleggeert-Lankamp CL, Peul WC: Interspinous process device versus standard conventional surgical decompression for lumbar spinal stenosis: randomized controlled trial. BMJ (Clinical research ed.) 347, f6415 (2013). doi:10.1136/bmj.f6415

26. Kim CH, Chung CK, Choi Y, Shin H, Woo JW, Kim SM, Lee HJ: The usefulness of a mobile device-based system for patient-reported outcomes in a spine outpatient clinic. The spine journal: official journal of the North American Spine Society 16(7), 843–850 (2016). doi:10.1016/j.spinee. 2016.02.048 [PubMed: 26961200]

27. Marcano Belisario JS, Jamsek J, Huckvale K, O'Donoghue J, Morrison CP, Car J: Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. The Cochrane database of systematic reviews(7), Mr000042 (2015). doi: 10.1002/14651858.MR000042.pub2

28. Zanoli G: Outcome assessment in lumbar spine surgery. Acta orthopaedica. Supplementum 76(318), 5–47 (2005).

29. Bouras T, Zairi F, Loufardaki M, Triffaux M, Stranjalis G: Which functional outcome parameters correlate better with elderly patients' satisfaction after non-fusion lumbar spine surgery? Journal of neurosurgical sciences (2017). doi:10.23736/s0390-5616.17.03977-7

30. Insel TR: Digital Phenotyping: Technology for a New Science of Behavior. Jama 318(13), 1215–1216 (2017). doi:10.1001/jama.2017.11295 [PubMed: 28973224]

31. Demir S, Dulgeroglu D, Cakci A: Effects of dynamic lumbar stabilization exercises following lumbar microdiscectomy on pain, mobility and return to work. Randomized controlled trial. European journal of physical and rehabilitation medicine 50(6), 627–640 (2014). [PubMed: 25201615]

32. Elsworth C, Dawes H, Winward C, Howells K, Collett J, Dennis A, Sackley C, Wade D: Pedometer step counts in individuals with neurological conditions. Clinical rehabilitation 23(2), 171–175 (2009). doi:10.1177/0269215508098895 [PubMed: 19164404]

33. Foley S, Quinn S, Jones G: Pedometer determined ambulatory activity and bone mass: a population-based longitudinal study in older adults. Osteoporosis international: a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA 21(11), 1809–1816 (2010). doi:10.1007/s00198-009-1137-1

34. Ammendolia C, Cote P, Rampersaud YR, Southerst D, Budgell B, Bombardier C, Hawker G: The boot camp program for lumbar spinal stenosis: a protocol for a randomized controlled trial. Chiropractic & manual therapies 24, 25 (2016). doi:10.1186/s12998-016-0106-y [PubMed: 27433335]

35. House JS, Landis KR, Umberson D: Social relationships and health. Science (New York, N.Y.) 241(4865), 540–545 (1988).

36. Dhand A, Luke DA, Lang CE, Lee JM: Social networks and neurological illness. Nature reviews. Neurology 12(10), 605–612 (2016). doi:10.1038/nrneurol.2016.119 [PubMed: 27615420]

37. Nagayoshi M, Everson-Rose SA, Iso H, Mosley TH Jr., Rose KM, Lutsey PL: Social network, social support, and risk of incident stroke: Atherosclerosis Risk in Communities study. Stroke 45(10), 2868–2873 (2014). doi:10.1161/strokeaha.114.005815 [PubMed: 25139878]

38. Steptoe A, Shankar A, Demakakos P, Wardle J: Social isolation, loneliness, and all-cause mortality in older men and women. Proceedings of the National Academy of Sciences of the United States of America 110(15), 5797–5801 (2013). doi:10.1073/pnas.1219686110 [PubMed: 23530191]

39. Bowling A: Mode of questionnaire administration can have serious effects on data quality. Journal of public health (Oxford, England) 27(3), 281–291 (2005). doi:10.1093/pubmed/fdi031

40. Johnson AM, Copas AJ, Erens B, Mandalia S, Fenton K, Korovessis C, Wellings K, Field J: Effect of computer-assisted self-interviews on reporting of sexual HIV risk behaviours in a general population sample: a methodological experiment. AIDS (London, England) 15(1), 111–115 (2001).
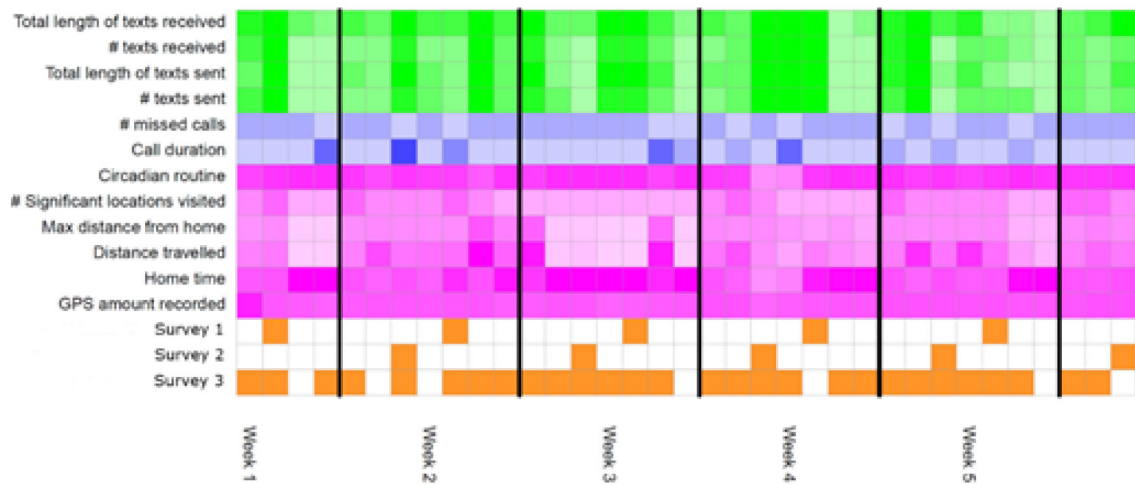
Figure 1. Visual display of daily summary statistics collected for a patient.

Surveys are represented in orange. Measures of mobility obtained from smartphone GPS are shown in purple. Measures from call logs are in blue and measures from text logs are in green. Measures are calculated from raw smartphone data on a daily basis. Only a smaller representative sample of the total set of measures both collected and calculated are displayed. Darker color represents a larger quantity/amount of the measure on a given day. Each column represents one day of data collection, with vertical black lines representing division between weeks.
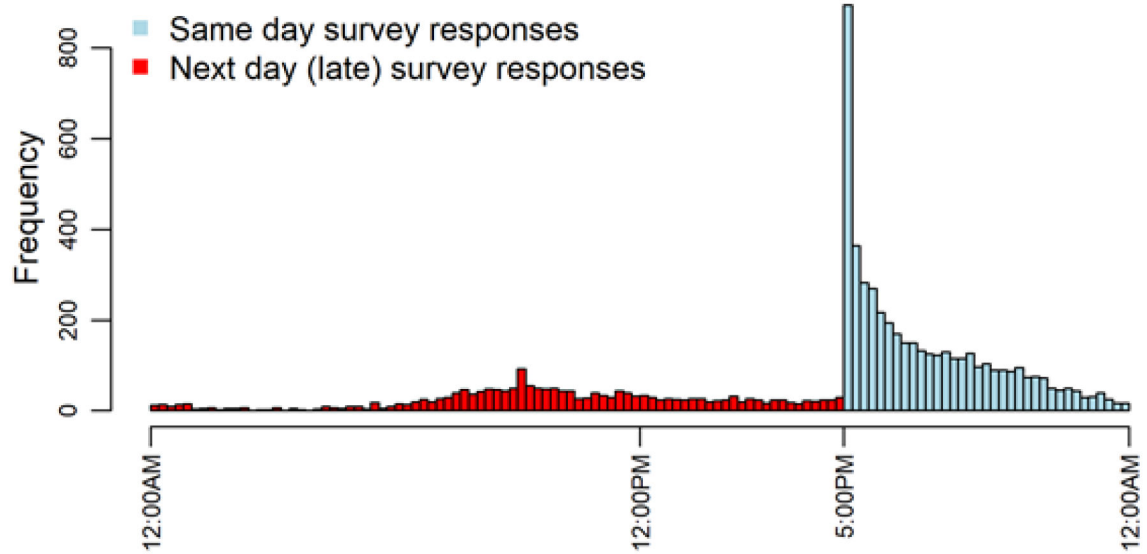
Figure 2. Response times to daily smartphone pain score survey.
Daily response times are combined across all patients in the cohort. Patients are prompted with a survey at 5:00PM each day where they are to report their pain from over the course of that day. Late responses submitted by the patient on the next day were excluded from the analysis. 70.7% of all survey responses were submitted on the same day the patient was prompted. Across all patients, 43% of these daily surveys were completed, late or on time. On average patients completed at least one survey per week 73.2% of the time.
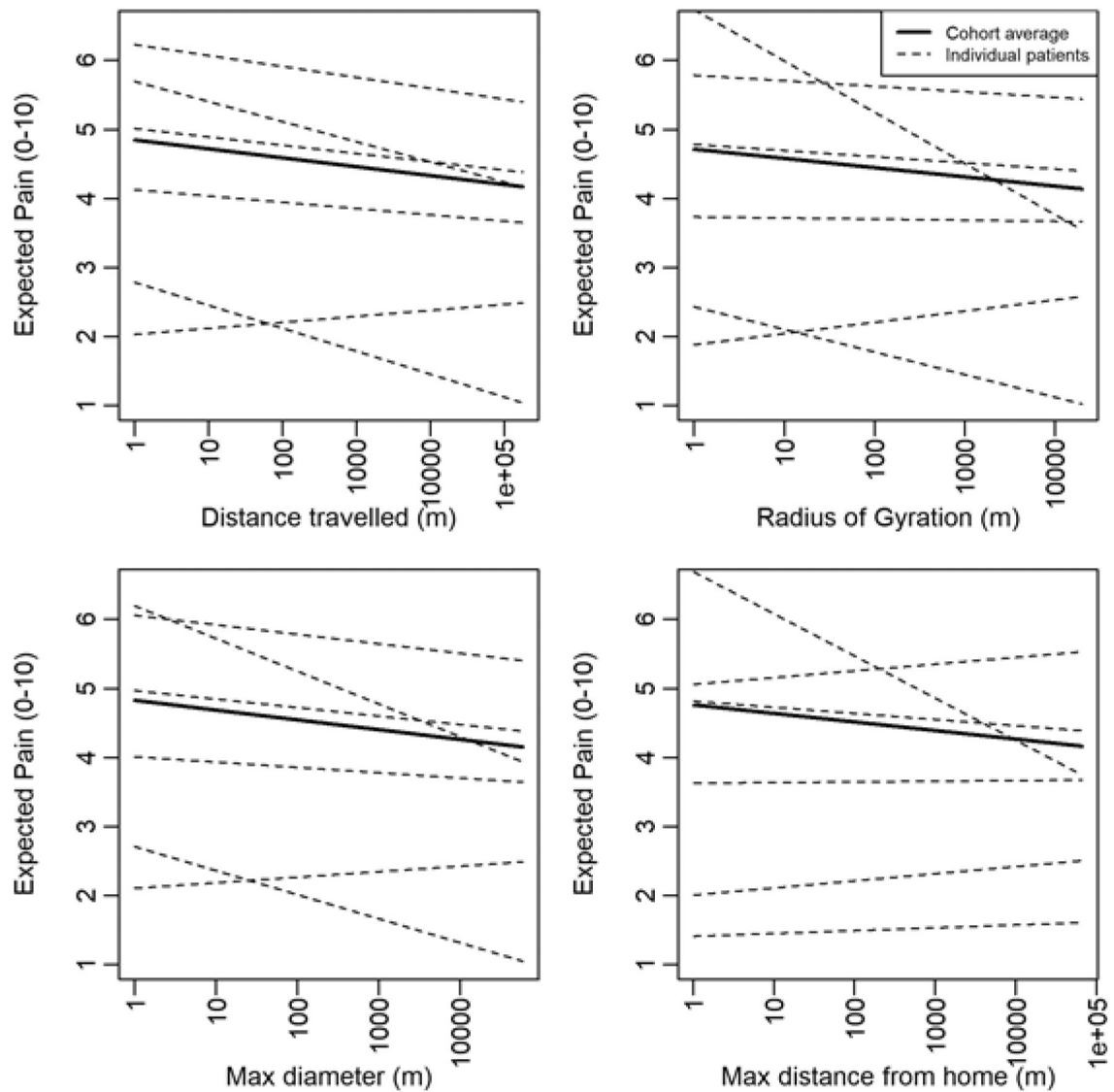
Figure 3. Patient-specific relationships between self-report of pain and mobility on the same day.
The six patients displayed are selected arbitrarily for demonstration purposes. The dotted lines represent the expected relationship between various mobility metrics and pain for individual patients, whereas the solid line is the average relationship across the cohort. These lines, representing the expected, or average, relationship between pain and mobility, result from fitting a linear mixed model with a random intercept and slope for each patient. A different model is fit for each mobility metric. The x-axis is on logarithmic scale.
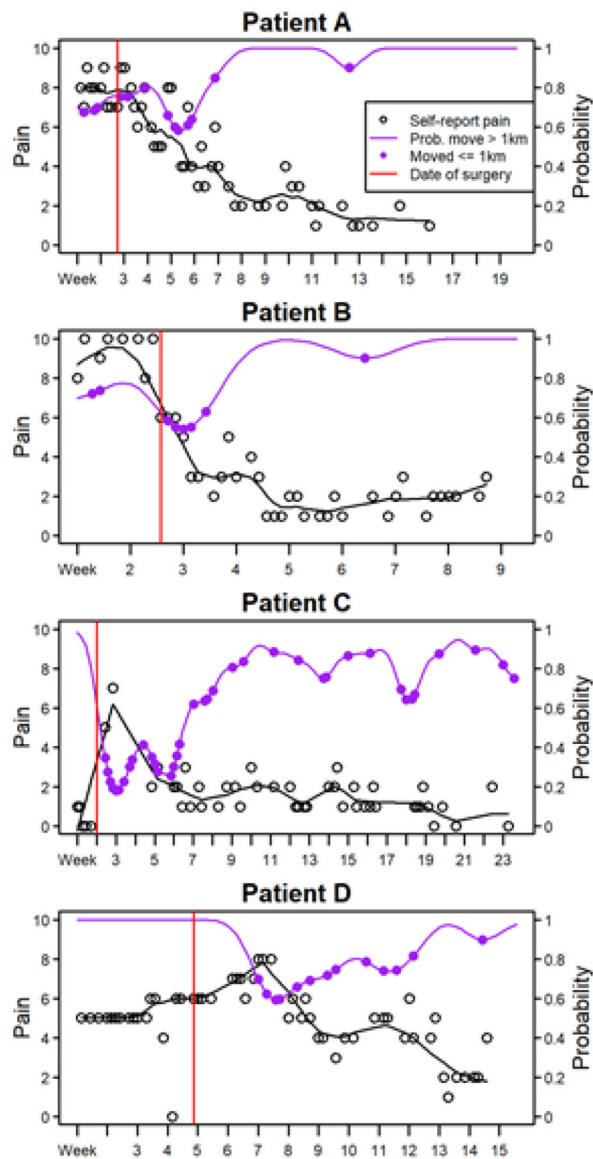
Figure 4. Patient pain and mobility trajectories over time.

The trajectories of four patients are displayed over the course of their recovery after surgery. X-axis tick marks represent weeks, and each point represents a day. The vertical red line is the date of surgery. The black line is the smoothed pain trajectory over time. Days where a patient travels less than 1km are highlighted with a purple point, with the patient traveling more than 1km on all other days. Some patients stopped taking their smartphone surveys after recovery despite continuing to collect smartphone GPS location and mobility data passively.

Table 1.

Participant demographic information and digital phenotyping data.

Demographics	Total (n=105)
Male (no., %)	48 (45.7)
Age (mean, st. dev.)	52.0 (14.0)
Surgery (no., %)	55 (52.4)
Site of Disease (no., %)	
Cervical	35 (33.3)
Thoracic	6 (5.7)
Lumbar	64 (60.1)
Passive Data Collection	
GPS Days (mean, st. dev.)	82.5 (68.4)
Active Data Collection	
Daily Survey Response Rate (mean, (25%,75%) quantiles)	43.4 (23.2,69.8)
Weekly Survey Response Rate (mean, (25%,75%) quantiles)	73.2 (50.6,100.0)
Change in Pain Score	
All patients (mean start/end score)	−1.3 (4.96 to 3.66)
Digital Phenotypes[*] Daily (median, (25%, 75%) quantiles)	
Time spent at home (minutes)	927.1 (623.9, 1242.8)
Distance travelled (meters)	51989.1 (18691.9, 97833.3)
Radius of gyration (meters)	3218.8 (846.8, 9881.9)
Maximum diameter (meters)	13828.1 (4867.5, 29382)
Maximum distance from home (meters)	13296.5 (4612.7, 29285.7)
Number of significant locations visited	2 (1, 3)
Average flight length (meters)	236.0 (160.1, 334.9)
Std. dev. of flight length (meters)	296.4 (173.3, 487.3)
Average flight duration (seconds)	42.7 (32.0, 74.3)
Std. dev. of flight duration (seconds)	98.2 (56.9, 257.1)
Fraction of the day not moving	0.88 (0.79, 0.94)
Significant location entropy	0 (0, 0.28)
Missing GPS data (minutes)	1349.7 (1323.7, 1379.2)
Circadian routine (0-low, 1-high)	0.59 (0.42, 0.71)
Week end/day stratified circadian routine	0.61 (0.44, 0.73)
Number of outgoing texts	4 (0, 14)
Total outgoing text length (characters)	165 (0, 756)
Texting outdegree	1 (0, 3)
Number of incoming texts	4 (0, 13)
Total incoming text length (characters)	188 (0, 689)
Texting indegree	2 (0, 4)

Demographics	Total (n=105)
Text reciprocity	4 (0, 15)
Text responsiveness	0.04 (0, 0.28)
Number of outgoing calls	1 (0, 4)
Total outgoing call lengths (seconds)	72 (0, 561)
Call outdegree	1 (0, 3)
Number of incoming calls	1 (0, 3)
Total incoming call lengths (seconds)	75 (0, 619)
Call indegree	1 (0, 2)
Call reciprocity	0 (0, 2)
Call responsiveness	0 (0, 0.44)

[*] For detailed descriptions and definitions of the mobility and sociability digital phenotypes, see Barnett and Onnela (2016).

Table 2.

Daily smartphone mobility and sociability features collected by digital phenotyping.

Data summary/feature	Collection type	Units
Pain	Active (survey)	0 (low) - 10 (high)
Time spent at home	Passive	Minutes ($\log_{10}$ scale)
Distance travelled	Passive	Meters ($\log_{10}$ scale)
Radius of Gyration	Passive	Meters ($\log_{10}$ scale)
Maximum diameter	Passive	Meters ($\log_{10}$ scale)
Maximum distance from home	Passive	Meters ($\log_{10}$ scale)
Number of significant locations visited	Passive	None
Avg. flight length	Passive	Meters ($\log_{10}$ scale)
Std. of flight length	Passive	Meters ($\log_{10}$ scale)
Avg. flight duration	Passive	Seconds ($\log_{10}$ scale)
Std. of flight duration	Passive	Seconds ($\log_{10}$ scale)
Fraction of time not moving	Passive	None
Significant location entropy	Passive	None
Circadian routine	Passive	None
Circadian routine (weekend/day stratified)	Passive	None
Number of outgoing texts	Passive	None
Cumulative length of outgoing texts	Passive	Characters ($\log_{10}$ scale)
Number of people texts were sent to	Passive	None
Number of incoming texts	Passive	None
Cumulative length of incoming texts	Passive	Characters ($\log_{10}$ scale)
Number of people texts were received from	Passive	None
Text reciprocity	Passive	None
Text responsiveness	Passive	None
Number of outgoing calls	Passive	None
Cumulative length of outgoing calls	Passive	Seconds ($\log_{10}$ scale)
Number of people calls were made to	Passive	None
Number of incoming calls	Passive	None
Cumulative length of incoming calls	Passive	Seconds ($\log_{10}$ scale)
Number of people calls were received from	Passive	None
Call reciprocity	Passive	None
Call responsiveness	Passive	None

Table 3.

Associations between daily passively collected variables and self-report of pain by linear mixed modeling each feature with a random slope and intercept for each participant.

Feature	Coef	Std. Err.	P-value
Average Flight Length	−0.226	0.068	0.002
Maximum Diameter	−0.142	0.047	0.004
Distance	−0.129	0.043	0.004
Radius of Gyration	−0.134	0.051	0.011
Std. dev. Flight Duration	−0.124	0.045	0.013
Std. dev. Flight Length	−0.174	0.071	0.020
Maximum Distance From Home	−0.124	0.055	0.029
Number of Significant Locations Visited	−0.740	0.408	0.077
Average Flight Duration	−0.108	0.062	0.089
Significant Location Entropy	−0.920	0.706	0.225
Text Responsiveness	0.576	0.456	0.247
Text Out-degree	0.284	0.256	0.320
Physical Circadian Routine (weekend/day stratified)	−0.632	0.737	0.408
Call In-degree	0.189	0.250	0.458
Text In-degree	0.223	0.257	0.474
Outgoing Texts	0.083	0.133	0.538
Call Responsiveness	0.255	0.435	0.568
Time Spent at Home	0.023	0.045	0.651
Physical Circadian Routine	−0.341	0.800	0.676
Number of Incoming Calls	0.084	0.213	0.694
Call Out-degree	−0.074	0.195	0.705
Call Reciprocity	−0.105	0.282	0.714
Number of Outgoing Calls	−0.044	0.195	0.820
Fraction of Time Stationary	−0.290	1.383	0.836
Number of Incoming Texts	0.030	0.142	0.837
Text Reciprocity	0.018	0.132	0.892
Total Incoming Call Lengths	0.006	0.045	0.892
Total Outgoing Text Lengths	0.007	0.060	0.909
Total Outgoing Call Lengths	−0.003	0.039	0.935