



Received: 19 October 2015
Accepted: 15 December 2015
First Published: 19 January 2016

*Corresponding author: David Trafimow,
Department of Psychology, New Mexico
State University, MSC 3452, P.O. Box
30001, Las Cruces, NM 88003-8001,
USA
E-mail: dtrafimo@nmsu.edu

Reviewing editor:
Udo Schumacher, University Medical
Center Hamburg-Eppendorf, Germany

Additional information is available at
the end of the article

ORTHOPEDICS | RESEARCH ARTICLE

The shocking implications of Bayes' theorem for diagnosing herniated nucleus pulposus based on MRI scans

David Trafimow^{1*} and Jordan H. Trafimow²

Abstract: We obtain the input data for Bayes Theorem, and use the theorem to determine the probability of a patient having a lumbar HNP, given only a positive MRI. We also enumerate the potential consequences that the clinician must keep in mind when making the diagnosis of lumbar HNP. We used the theorem by Bayes, in conjunction with well-established results in the orthopedic literature, to calculate the probability of lumbar HNP given only a positive MRI finding. The necessary information provided by the orthopedic literature includes the prevalence of lumbar HNP, the probability of a positive MRI finding given that there is no HNP, and the probability of a positive MRI finding given that there is HNP. We found that the probability of lumbar HNP given only a positive MRI finding was 8%. The probability that there is no lumbar HNP, even when there is a positive MRI finding, is 92%. Clearly, MRI scans cannot be trusted as the sole source of diagnostic information

Subjects: Health Conditions; Medicine; Medicine, Dentistry, Nursing & Allied Health

Keywords: Bayes' theorem; herniated nucleus pulposus; MRI scans

1. Introduction

Herniated lumbar discs (HNP) are a large problem in the Western world. Andersson found that 1–3% of the population of Finland and Italy had HNP (Andersson, 1991). Researchers who have investigated other populations obtained similar findings (Borenstein, Wiesel, & Boden, 1995; Frymoyer, 1988; Lawrence et al., 2008). This condition is one of much pain and suffering for the afflicted patient, to say nothing of the financial cost.

ABOUT THE AUTHORS

David Trafimow is a distinguished achievement professor of psychology at New Mexico State University, a fellow of the Association for Psychological Science, executive editor of the *Journal of General Psychology*, and also for *Basic and Applied Social Psychology*. He received his PhD in psychology from the University of Illinois at Urbana-Champaign in 1993. His current research interests include attribution, attitudes, cross-cultural research, ethics, morality, methodology, and potential performance theory.

Jordan H. Trafimow is an orthopedist with experience as a practicing doctor, as a medical researcher, and in administrative roles. He is currently retired from active surgery and is concentrating on medical research.

PUBLIC INTEREST STATEMENT

The article, “The shocking implications of Bayes' theorem for diagnosing herniated nucleus pulposus based on MRI scans” shows that because of the low base rate for the disorder, and the substantial false alarm rate, the probability that a person has the disorder given that an MRI scan has diagnosed it, is only 8%. Thus, the probability of not having herniated nucleus pulposus (commonly called a “slipped disk”) is 92% when an MRI scan is positive. The importance of a proper examination is emphasized, as is the demonstration of the lack of validity of MRI scans for these sorts of diagnoses.

If surgery is needed, the costs increase considerably. This is worthwhile if the chances of helping the patient are good. If the diagnosis is wrong, the patient has taken the risks of surgery and suffered the pain and financial cost for no good reason. The real tragedy is when a patient develops a “failed back syndrome” from the surgery when the diagnosis was wrong and the patient should never have had the surgery.

Making the diagnosis of HNP is not always easy. Doctors use three types of information in making the diagnosis of HNP: History and physical exam (H & P), Electrodiagnostic studies, usually EMG/NCV, and Radiology, usually MRI (Cho, Ferrante, Levin, Harmon, & So, 2010). Components of and problems with each are summarized in the following sections.

2. H&P

We first consider the contents of the exam; different sources have somewhat different thoughts.

- The New York Mid and Low Back Injury Medical Treatment Guidelines First Edition 30 June 2010 require details of muscle weakness (The New York mid & low back injury medical treatment guidelines, 2010).
- The American College of Occupational and Environmental Medicine guidelines (the official guidelines for the state of California) require history of pain and paresthesia location, and examination of the muscles innervated by each nerve root (Glass, 2004).
- The Official Disability Guidelines (the official guidelines for the state of Texas) require exam of Quadriceps, Tibialis Anterior, toe and ankle plantar flexors, straight leg raising (SLR) and crossed straight leg raising, and reflex exams (Denniston, 2009).
- Wheeler et al. require that exam contains SLR and crossed SLR, Extensor Hallucis Longus, Tibialis Anterior, Triceps Surae, ankle jerk, and location of sensory decrease (Wheeler, Wipf, Staiger, et al., 2010).
- Atlas and Deyo required hip flexion, knee extension (L3 and L4), Tibialis Anterior, Extensor Hallucis Longus, and Triceps Surae (Atlas & Deyo, 2001).
- Hakelius and Hindmarsh required knee jerk as the only finding that mattered for L3-4 HNP (Hakelius & Hindmarsh, 1972).
- Weber required radicular pain, positive SLR, or muscle weakness (Weber, 1983).

Multiple nerve roots might innervate a single muscle. If that muscle is weak, then at least one of the nerve roots is inflamed. In the absence of other evidence, an MRI is indicated to locate the presumed HNP.

All nerve roots innervate multiple muscles. Consequently, all of the muscles must be examined because any one of them might be weak. If one of them actually is weak, an MRI is indicated.

Some patients have radicular pain. Some doctors might be tempted to conclude that this indicates HNP and so an MRI is indicated. However, it is common that such pain indicates pathology different from HNP; one example is a trigger point, which causes pain radiation down a leg in an L5 or S1 distribution.

In summary, we recommend the examination of all the muscles or joint motions that are innervated by L4, L5, and S1. According to Chou et al., at least 90% of HNP occurs at L4-5 and L5-S1 (Chou et al., 2007). As noted above, weakness of a muscle innervated by multiple roots is still evidence of nerve root irritation.

The following problems are standard in the evaluation by H & P:

- (1) Crossed SLR is 90% specific for HNP but not very sensitive whereas SLR is more sensitive but less specific (Wheeler et al., 2010).
- (2) Muscles may be weak because of pain or disuse or other reasons, rather than HNP.
- (3) Not all the muscles innervated by an affected nerve root are weak in every case.
- (4) Anatomic variations happen in about 3% of cases, resulting in pre- or post-fixed lumbar plexus (e.g. the L5 root carrying the fibers normally considered L4 or S1) (Hollinshead, 1958).
- (5) Authorities do not agree on which nerve root(s) supply certain muscles. Andersson and Deyo stated that the Peroneals are innervated by S1 and ankle dorsiflexors are innervated by L5; Extensor Hallucis Longus (EHL) weakness is found more often in patients with L5-S1 HNP than L4-5 HNP (Andersson & Deyo, 1996). Wheeler et al. wrote that the EHL and Tibialis Anterior get innervation from L5 (Wheeler et al., 2010). Atlas and Deyo said that L4 innervates the Tibialis Anterior and L5 innervates the Extensor Hallucis Longus, Quadriceps innervation is from L3 and L4 (Atlas & Deyo, 2001). Finally, Hollinshead wrote that L5 and S1 innervate the Gluteus Maximus, and the Gluteus Medius's and Minimus's nerve supply is from L4, L5, and S1. Also according to Hollinshead, L5 and S1 supply the Extensor Hallucis Longus; L4-S1 supplies the Peroneals (Hollinshead, 1958).

3. EMG/NCV

EMG/NCV is useful in situations where one cannot trust the muscle exam. However, they will be normal if the nerve root is irritated enough to cause pain, but not enough to show injury electrically. Glantz and Haldeman discussed further problems with the EMG: (Glantz & Haldeman, 1991)

- (1) Weak muscles may have normal EMG.
- (2) NCV is usually normal in radiculopathy.
- (3) Even with sensory loss, sensory EMG/NCV is usually normal
- (4) Completely reinnervated muscles will be normal by EMG.

Similarly, Cho et al. stated that electrodiagnostic studies do not make any independent contribution to the diagnosis of lumbosacral radiculopathy (Cho et al., 2010).

4. MRI

Several authorities have come to conclusions about the validity of MRI findings.

- (1) Boden et al. found that the MRI is 30% falsely positive (Boden, Davis, Dina, et al., 1990). Jensen et al. obtained similar results (Jensen et al., 1994). If the MRI does not show a herniated disc, it is accepted that there is no herniated disc. Of course there may be positive findings from other diagnoses, such as chemical radiculitis (Friedman & Goldner, 1983; Marshall, Trethewie, & Curtain, 1977).
- (2) Radiologists often report the radiologic findings, but they say that the diagnosis should be made with the clinical findings considered. In other words, if the clinical findings lead to a diagnosis of nerve root problem, a positive MRI finding shows HNP. But if not, a positive MRI finding does not show HNP.
- (3) There is no unanimity about the meanings of various words used to describe an abnormal bulge in a disk. A survey by NASS found that "extruded" had clear meaning, but the members could not clearly distinguish between bulging, slipped, and herniated (Fardon & Milette, 2001).

The foregoing problems with MRI validity suggest that its use, in the absence of clinical evidence, might not increase clinical results. Chou, Fu, Carrino, and Deyo (2009) tested this possibility in a

systematic review and meta-analysis and found that short- or long-term outcomes did not vary significantly between those patients who received MRI and those with conventional care only. They concluded that clinicians should not get routine imaging (MRI or CT) unless there are clinical findings suggesting a serious underlying condition (Chou et al., 2009). This conclusion is consistent with Vucetic, Astrand, Guentner, and Svensson (1999) who stated: “Many people have asymptomatic herniations, and today supersensitive imaging is widely available. Thus the importance of clinical evaluation has increased, and most of the relevant information can be obtained by listening to the patient” (Vucetic et al., 1999). Staiger et al. (2010) agrees that because of the problems with H & P and EMG/NCV discussed earlier, many doctors place undue reliance on the MRI and consequently approximately ¼ of patients who obtained an MRI did so without indication (Staiger, Gatewood, Wipf, et al., 2010).

Because of imaging difficulties similar to the above, Charles Herndon, M.D. taught his residents in the 1960s that “IF TO OPERATE” is a clinical decision, but “WHERE TO OPERATE” requires imaging confirmation.” However, Rubinstein and Tuldar stated that “the clinician can accurately identify sciatica due to disk herniation” (Rubinstein & Tuldar, 2008). Wheeler et al. wrote that the only reason for MRI is worsening neurological deficit (Wheeler et al., 2010).

5. Hypothesis

In support of Dr. Herndon’s teaching, our hypothesis is “The probability of finding a herniated disc at surgery is very low if diagnosis is made only by virtue of a positive MRI finding, without positive clinical findings.”

6. Purpose

The purpose of this study is to prove the hypothesis by using the famous theorem by Bayes. The orthopedic literature provides the data to use in the theorem. Such proof would provide a strong argument in favor of not relying too heavily on only MRI findings.

7. Method

Bayes’ theorem is used in a wide variety of areas such as mathematics, statistics, engineering, physics, neuroscience, and others. It takes on a variety of forms but the form that is convenient for our purposes is given below as Equation (1). We use Equation (1) to obtain the conditional probability of lumbar HNP given a positive finding from an MRI scan [$P(HNP|F)$]. A conditional probability is the probability of one thing given that something else is so, and the vertical line in $P(HNP|F)$ symbolizes “given that.” The variables on the other side of the equals sign are the prevalence of lumbar HNP [$P(HNP)$], the probability of a positive MRI finding given that there is HNP [$P(F|HNP)$], and the probability of a positive MRI finding when there is no HNP [$P(F|\sim HNP)$]. The tilde is the symbol for negation.

$$P(HNP|F) = \frac{P(HNP)P(F|HNP)}{P(HNP)P(F|HNP) + P(F|\sim HNP)[1 - P(HNP)]} \quad (1)$$

To use Bayes’ theorem, it is necessary to have three items of information. First, we need the probability of a positive MRI finding given that there is no HNP; this is 30% (Jensen et al., 1994; Marshall et al., 1977). Second, we need the probability of a positive MRI finding given that there is HNP. According to Boos et al., this probability is 80.4% (Boos et al., 1995). Third, we need the prevalence of lumbar HNP in the population, which we already have seen is between 1 and 3% (Andersson, 1991; Borenstein et al., 1995; Frymoyer, 1988; Lawrence et al., 2008). The usual statistical term for prevalence is “base rate” and so we will use the latter term hereafter. To be conservative, we will use 3% for the base rate, as the findings to be presented would be even more extreme if we used a lower number.

8. Results

After substituting the numbers provided in the foregoing paragraph into the appropriate places in Equation (1), we find that the probability of lumbar HNP given a positive MRI finding is .0765, or under 8%. Obviously, then, in approximately 92% of the cases where lumbar HNP is diagnosed based solely on a positive MRI finding, there actually is no lumbar HNP.

Given the low base rate for lumbar HNP (3%), the finding that the probability of lumbar HNP given a positive MRI result is less than 8% is not mathematically surprising to researchers who are familiar with Bayesian analyses. However, to the many doctors who confidently diagnose lumbar HNP based on positive MRI results, and who are unfamiliar with Bayesian analyses, our finding should indeed be revealing.

9. Discussion

Intuition might suggest that the 30% false positive rate implies that the MRI is correct 70% of the time. But this intuition contradicts our finding and the driving force behind the contradiction is the low base rate for lumbar HNP. It is because of this low base rate that when there is a positive MRI finding, the conclusion that there is lumbar HNP has a 92% probability of being wrong. The fact that intuition and hard mathematics are in contradiction demonstrates the value of depending on hard mathematics rather than on intuition. This contradiction also demonstrates the value of considering base rates. From the perspective of the practicing doctor, the findings cast doubt on the validity of the MRI in diagnosing lumbar HNP. In essence, it means that the doctor should not order an MRI to diagnose lumbar HNP without having clinical evidence.

One way to conceptualize positive clinical findings is that the positive findings increase the base rate of lumbar HNP. Put another way, the percentage of people with lumbar HNP likely is far greater in patients with positive clinical indications than in the general population. With a substantial increase in the base rate, the combination of a complete clinical examination and an MRI is more valid than an MRI alone for diagnosing the presence of HNP. We hasten to add that once the presence of HNP is diagnosed, an MRI can be invaluable for diagnosing where it is.

Consider a hypothetical example where the clinical findings are positive and so the base rate probability that a particular patient has lumbar HNP is 70% as opposed to 3% in the general population. In that case, if the MRI finding is also positive, applying Equation (1) indicates that the conditional probability of HNP given the positive MRI finding is 86% and so the probability of wrongly drawing the conclusion is 14%. Obviously, it is better to be wrong 14% of the time than to be wrong 92% of the time, thereby illustrating the potential importance of a complete clinical examination.

The foregoing hypothetical example suggests an important direction for future research. If doctors knew the base rate of HNP given a single positive clinical finding, two positive clinical findings, and so on, this information could be used in Equation (1) to calculate, for each patient, the conditional probability of HNP given both the clinical picture and the MRI finding. If different clinical findings are of unequal diagnostic values, this also could be figured into Equation (1). Thus, we strongly suggest that researchers collect the requisite data to obtain these probabilities. More generally, a recent introduction and tutorial on the use of Bayesian analyses in medicine suggests a variety of ways to make important medical advances through Bayesian methods (Trafimow, 2015).

It is unfortunate when a patient undergoes indicated surgery that, though justified, nevertheless causes complications such as epidural scarring or “failed back syndrome” that condemn him or her to serious lifelong problems. But if the surgeon, based on a positive MRI finding in the absence of a proper clinical examination, operates when it is not justified, the complications are an unmitigated tragedy. We hope that orthopedists will take the present Bayesian lesson to heart, perform proper clinical examinations, and thereby reduce the incidence of unjustified operations in the future.

Funding

The authors received no direct funding for this research.

Competing interests

The authors declare no competing interest.

Author details

David Trafimow¹

E-mail: dtrafimo@nmsu.edu

Jordan H. Trafimow²

E-mail: jhtrafimow@comcast.net

ORCID ID: <http://orcid.org/0000-0003-0253-0534>

¹ Department of Psychology, New Mexico State University, MSC 3452, P.O. Box 30001, Las Cruces, NM 88003-8001, USA.

² The ARK, Chicago, USA.

Citation information

Cite this article as: The shocking implications of Bayes' theorem for diagnosing herniated nucleus pulposus based on MRI scans, David Trafimow & Jordan H. Trafimow, *Cogent Medicine* (2016), 3: 1133270.

References

- Andersson, G. B. (1991). Epidemiology of spinal disorders. In J. Frymoyer (Ed.), *The adult spine: Principles and practice* (pp. 107–146). New York, NY: Raven Press.
- Andersson, G. B., & Deyo, R. A. (1996). History and physical examination in patients with herniated lumbar discs. *Spine*, 21, 105–185. <http://dx.doi.org/10.1097/00007632-199612151-00003>
- Atlas, S. J., & Deyo, R. A. (2001). Evaluating and managing acute low back pain in the primary care setting. *Journal of General Internal Medicine*, 16, 120–131. <http://dx.doi.org/10.1111/j.1525-1497.2001.91141.x>
- Boden, S. D., Davis, D. O., Dina, T. S., et al. (1990). Lumbar spine in asymptomatic subjects: A prospective investigation. *The Journal of Bone & Joint Surgery*, 72, 403–407.
- Boos, N., Rieder, R., Schade, V., Spratt, K. F., Semmer, N., & Aebi, M. (1995). The diagnostic accuracy of magnetic resonance imaging, work perception, and psychosocial factors in identifying symptomatic disc herniations. *Spine*, 20, 2613–2625. <http://dx.doi.org/10.1097/00007632-199512150-00002>
- Borenstein, D. G., Wiesel, D. G., & Boden, S. D. (1995). *Low back pain-medical diagnosis and comprehensive management* (2nd ed.). Philadelphia, PA: WB Saunders.
- Cho, S. C., Ferrante, M. A., Levin, K. H., Harmon, R. L., & So, Y. T. (2010). Utility of electrodiagnostic testing in evaluating patients with lumbosacral radiculopathy: An evidence-based review. *Muscle & Nerve*, 42, 276–282.
- Chou, R., Fu, R., Carrino, J., & Deyo, R. A. (2009). Imaging strategies for low-back pain: Systematic review and meta-analysis. *The Lancet*, 373, 463–472. [http://dx.doi.org/10.1016/S0140-6736\(09\)60172-0](http://dx.doi.org/10.1016/S0140-6736(09)60172-0)
- Chou, R., Qaseem, A., Snow, V., Casey, D., Cross, J. T., Shekelle, P., & Owens, D. K. (2007). Diagnosis and treatment of low back pain: A joint clinical practice guideline from the American college of physicians and the American pain society. *Annals of Internal Medicine*, 147, 478–491. <http://dx.doi.org/10.7326/0003-4819-147-7-200710020-00006>
- Denniston, P. L. (Ed.). (2009). *Official disability guidelines*. Encinitas, CA: Work Loss Data Institute.
- Fardon, D. F., & Milette, P. D. (2001). Nomenclature and classification of lumbar disc pathology. *Spine*, 26, E83–E113.
- Friedman, J. I., & Goldner, M. Z. (1983). Chemical radiculitis: A clinical, physiological, and immunological study. *Clinical Research*, 31, 650 A.
- Frymoyer, J. W. (1988). Epidemiology. In J. W. Frymoyer & S. L. Gordon (Eds.), *New perspectives on low back pain* (pp. 19–33). (Symposium, Workshop, Airlie). Chicago, IL: American Academy of Orthopedic Surgeons.
- Glantz, R. H., & Haldeman, S. (1991). Other diagnostic studies: Electrodiagnosis. In J. Frymoyer (Ed.), *The adult spine: Principles and practice* (pp. 541–548). New York, NY: Raven Press.
- Glass, L. S. (Ed.). (2004). *Occupational medicine practice guidelines evaluation and management of common health problems and functional recovery of workers* (2nd ed.). Elk Grove Village, IL: American College of Occupational and Environmental Medicine.
- Hakelius, A., & Hindmarsh, J. (1972). The comparative reliability of pre operative methods in lumbar disc surgery. *Acta Orthopaedica Scandinavica*, 43, 243–248.
- Hollinshead, H. (1958). *Anatomy for surgeons Part III. The back and limbs* (pp. 676–679). New York, NY: Hoeber-Harper.
- Jensen, M. C., Brant-Zawadzki, M. N., Obuchowski, N., Modic, M. T., Malkasian, D., & Ross, J. S. (1994). Magnetic resonance imaging of the lumbar spine in people without back pain. *New England Journal of Medicine*, 331, 69–73. <http://dx.doi.org/10.1056/NEJM199407143310201>
- Lawrence, R. C., Feseon, D. T., Helmick, C. G., Arnold, L. M., Choi, H., Deyo, R. A., ... Wolfe, F. (2008). Estimates of the prevalence of arthritis and other rheumatic conditions in the United States Part II. *Arthritis & Rheumatism*, 58, 26–35.
- Marshall, L. L., Trethewie, E. R., & Curtain, C. C. (1977). Chemical radiculitis. *Clinical Orthopaedics and Related Research*, 129, 61–67. <http://dx.doi.org/10.1097/00003086-197711000-00006>
- Rubinstein, S. M., & Tuldar, M. (2008). A best-evidence review of diagnostic procedures for neck and low-back pain. *Clinical Rheumatology*, 22, 471–482.
- Staiger, T. O., Gatewood, M., Wipf, J. E., et al. (2010). Diagnostic testing for low back pain. In S. J. Atlas (Ed.), *UpToDate* (Vol. 18). Alphenan der Rhein: Wolter Kluver.
- The New York mid and low back injury medical treatment guidelines. (2010, June 30). (1st ed.).
- Trafimow, D. (2015). The benefits of applying Bayes' theorem in medicine. *American Research Journal of Humanities and Social Sciences*, 1, 14–23.
- Vucetic, N., Astrand, P., Guentner, P., & Svensson, O. (1999). Diagnosis and prognosis in lumbar disc herniation. *Clinical Orthopaedics and Related Research*, 361, 116–122. <http://dx.doi.org/10.1097/00003086-199904000-00016>
- Weber, H. (1983). Lumbar disc herniation a controlled, prospective study with ten years of observation. *Spine*, 8, 131–140. <http://dx.doi.org/10.1097/00007632-198303000-00003>
- Wheeler, S. G., Wipf, J. E., Staiger, T. O., & Deyo, R. A. (2010). Approach to the diagnosis and evaluation of low back pain in adults. In S. J. Atlas (Ed.), *UpToDate*. (Vol. 18). Alphenan der Rhein: Wolter Kluver.

PREPRINT ONLY



© 2016 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.



Cogent Medicine (ISSN: 2331-205X) is published by Cogent OA, part of Taylor & Francis Group.

Publishing with Cogent OA ensures:

- Immediate, universal access to your article on publication
- High visibility and discoverability via the Cogent OA website as well as Taylor & Francis Online
- Download and citation statistics for your article
- Rapid online publication
- Input from, and dialog with, expert editors and editorial boards
- Retention of full copyright of your article
- Guaranteed legacy preservation of your article
- Discounts and waivers for authors in developing regions

Submit your manuscript to a Cogent OA journal at www.CogentOA.com

