How Accurate Are Patients at Diagnosing the Cause of Their Knee Pain With the Help of a Web-based Symptom Checker?

Leslie J. Bisson,*† MD, Jorden T. Komm,† MD, Geoffrey A. Bernas,† MD, Marc S. Fineberg,† MD, John M. Marzo,† MD, Michael A. Rauh,† MD, Robert J. Smolinski,† MD, and William M. Wind,† MD

Investigation performed at University Orthopaedics and Sports Medicine outpatient locations, The State University of New York at Buffalo, Buffalo, New York, USA

Background: Researching medical information is the third most popular activity online, and there are a variety of web-based symptom checker programs available.

Purpose: This study evaluated a patient's ability to self-diagnose their knee pain from a list of possible diagnoses supplied by an accurate symptom checker.

Study Design: Cohort study (diagnosis); Level of evidence, 2.

Methods: All patients older than 18 years who presented to the office of 7 different fellowship-trained sports medicine surgeons over an 8-month period with a complaint of knee pain were asked to participate. A web-based symptom checker for knee pain was used; the program has a reported accuracy of 89%. The symptom checker generates a list of potential diagnoses after patients enter symptoms and links each diagnosis to informative content. After exploring the informative content, patients selected all diagnoses they felt could explain their symptoms. Each patient was later examined by a physician who was blinded to the differential generated by the program as well as the patient-selected diagnoses. A blinded third party compared the diagnoses generated by the program with those selected by the patient as well as the diagnoses determined by the physician. The level of matching between the patient-selected diagnoses and the physician's diagnoses determined the patient's ability to correctly diagnose their knee pain.

Results: There were 163 male and 165 female patients, with a mean age of 48 years (range, 18-76 years). The program generated a mean 6.6 diagnoses (range, 2-15) per patient. Each patient had a mean 1.7 physician diagnoses (range, 1-4). Patients selected a mean 2 diagnoses (range, 1-9). The patient-selected diagnosis matched the physician's diagnosis 58% of the time.

Conclusion: With the aid of an accurate symptom checker, patients were able to correctly identify the cause of their knee pain 58% of the time.

Keywords: knee; knee pain; diagnosis; symptom checker

Researching medical information has been reported as the third most popular activity online, and there are a variety of web-based symptom checker programs available to the patient.[4] A recent study has demonstrated that a web-based program can generate an accurate differential diagnosis in 89% of ambulatory patients with knee pain based solely on the history entered by the patient.[1] Despite the regularity of online symptom checking, we are not aware of any study that has evaluated a patient's ability to self-diagnose the cause of their knee pain with the assistance of a symptom checker. The purpose of this study was to determine patients' ability to select their diagnosis from a list of possible diagnoses supplied by an accurate symptom checker.

METHODS

Approval for this study was granted by our institutional review board, and consent was obtained. All patients older than 18 years who presented to the office of 7 different fellowship-trained sports medicine surgeons over an

*Address correspondence to Leslie J. Bisson, MD, School of Medicine and Biomedical Sciences, The State University of New York at Buffalo, 462 Grider Street, Buffalo, NY 14215, USA (email: ljbisson@buffalo.edu).

†School of Medicine and Biomedical Sciences, The State University of New York at Buffalo, Buffalo, New York, USA.

One or more of the authors has declared the following potential conflict of interest or source of funding: This study was funded by the Ralph and Mary Wilson Fund.

The Orthopaedic Journal of Sports Medicine, 4(2), 2325967116630286
DOI: 10.1177/2325967116630286

TABLE 1
Possible Diagnoses

Anterior cruciate ligament tear
Inflammatory arthritis
Iliotibial band friction syndrome
Lateral collateral ligament tear
Medial collateral ligament tear
Meniscal tear
Osgood-Schlatter disease
Osteoarthritis
Osteoarthritis exacerbation
Osteochondritis dissecans
Patellar arthritis
Patellar arthritis exacerbation
Patellar chondromalacia/patellofemoral syndrome
Patellar contusion/saphenous nerve contusion
Patellar instability
Patellar tendinitis
Patellar tendon rupture (partial or complete)
Plica syndrome
Popliteal cyst
Posterior cruciate ligament tear
Prepatellar bursitis
Quadriceps tendinitis
Quadriceps tendon tear (partial or complete)
Stress fracture
Trochlear chondromalacia
Rheumatoid arthritis

8-month period (June 2014–January 2015) with a complaint of knee pain were asked to participate. Exclusion criteria included patients who did not complete the program in its entirety, those younger than 18 years, patients who did not have a progress note supplying a physician's diagnosis or a clinical diagnosis was not recorded in the chart, and those who had a physician diagnosis that was not capable of being generated by the program.

We utilized a web-based symptom checker for knee pain that has been described previously, which generates a differential diagnosis following patient-entered symptoms and has a reported accuracy of 89%.[1] Briefly, after entry of symptoms, the program supplies a list of potential diagnoses that are selected from 26 possible diagnoses (Table 1). As in the original study,[1] for the purposes of analysis, the 26 diagnoses were consolidated to 21. First, osteoarthritis and osteoarthritis exacerbation were considered a single diagnosis. For example, if the clinical diagnosis was osteoarthritis and the program generated osteoarthritis exacerbation, this was considered an accurate match. The rationale for combining osteoarthritis and osteoarthritis exacerbation was because they are symptomatically the same diagnosis, with the only difference being a traumatic onset of pain in the case of osteoarthritis exacerbation. Furthermore, they are treated the same, that is, with nonsteroidal anti-inflammatory drugs (NSAIDs), physical therapy, injections, or some combination of these treatments. The same logic was used for patellar arthritis and patellar arthritis exacerbation. Additionally, 4 diagnoses generated by the program were grouped together as "patellofemoral pain" and considered 1 diagnosis. These included patellar chondromalacia/patellofemoral syndrome, patellar contusion/saphenous nerve contusion, plica syndrome, and trochlear chondromalacia. The rationale for combining these diagnoses was that the treatment approach to these 4 diagnoses is the same (rest, ice, NSAIDs, and physical therapy). The influence of combining them would therefore not be expected to change treatment.

The program was expanded for this study as follows: Each diagnosis was linked to informative content, including a description of the diagnosis, methods an individual would use to determine if this diagnosis is applicable to their knee pain, and a video describing physical examination findings pertinent to the diagnosis. The content is available by visiting the website at http://www.virtualkneedoc.com. The informative content was linked to sites of recognized authoritative bodies, including the American Academy of Orthopaedic Surgeons. The senior author (L.J.B.) created the additional narrative content included in the information. Also, directly under the list of diagnoses the program supplied in response to their symptoms, the patients were given a list of the remaining diagnoses that could be generated by the program, with each of those diagnoses also being linked to informative content. Patients participating in the study were asked to explore the content for each diagnosis supplied by the program, followed by review of the content for the remaining diagnoses that were not generated by the program for their particular symptoms. Finally, they were presented with a list of the diagnoses and asked to check all they felt could explain their symptoms. The program along with its informative content can be viewed at http://www.virtualkneedoc.com.

Within days of completing the program, each patient was examined by a board-certified, sports medicine fellowship–trained orthopaedic surgeon. The physician performed their workup. In almost all cases, the diagnosis was reached with the inclusion of radiographs, and in some cases diagnosis was reached with the aid of more advanced imaging. The physician was blinded to the differential diagnosis generated by the program as well as the diagnoses selected by the patient. Finally, a blinded third party compared the diagnoses generated by the program with that selected by the patient as well as the final diagnosis determined by the physician. The diagnosis provided by the physician was considered the correct diagnosis. The level of matching of the program-generated diagnoses and those supplied by the physician were used to calculate the sensitivity and specificity of the program. The level of matching between the patient-selected diagnoses and the physician's diagnoses determined the ability of the patient to correctly diagnose the cause of his or her knee pain.

The sensitivity and specificity of the program used in this study were calculated using Microsoft Excel 2010 (Microsoft Corp). Sensitivity was defined as the ability of the program to generate the physician's diagnosis as part of its differential. The sensitivity for each diagnosis was calculated as the total number of matches for that diagnosis divided by the number of times that diagnosis was given by a physician. The specificity of the program was defined as how often a given diagnosis produced by the program was indeed the correct diagnosis given by the physician. Specificity was calculated as total number of matches

TABLE 2
Summary for Each Diagnosis

Diagnosis	Patient Selected, n	Program Generated, n	Physician Diagnosed, n	Patient Matched Physician, n	Correct Patient Selection, %	Program Matched Physician, %
Anterior cruciate ligament tear	33	115	29	20	69	93
Inflammatory arthritis	22	130	0	0	—	—
Iliotibial band friction syndrome	16	48	4	3	75	75
Lateral collateral ligament tear	13	39	0	0	—	—
Medial collateral ligament tear	16	52	10	4	40	60
Meniscal tear	153	229	131	94	72	83
Osgood-Schlatter disease	10	73	2	1	50	100
Osteoarthritis/osteoarthritis exacerbation	144	328	157	97	62	100
Osteochondritis dissecans	0	3	0	0	—	—
Patellar arthritis/patellar arthritis exacerbation	80	298	100	41	41	95
Patellofemoral pain	81	277	82	35	43	91
Patellar instability	15	92	4	4	100	100
Patellar tendinitis	10	72	8	2	25	50
Patellar tendon rupture (partial or complete)	2	50	0	0	—	—
Popliteal cyst	18	65	4	4	100	100
Posterior cruciate ligament tear	5	18	3	2	67	67
Prepatellar bursitis	15	103	3	2	67	100
Quadriceps tendinitis	6	38	1	1	100	100
Quadriceps tendon tear (partial or complete)	5	60	2	2	100	100
Stress fracture	3	53	0	0	—	—
Rheumatoid arthritis	6	18	3	3	100	67
Total	653	2161	543	315	58	91

divided by the number of times a given diagnosis was produced by the program. Sensitivity and specificity of the patient-selected diagnoses was also calculated. The sensitivity of the patient-selected diagnosis was defined as the ability of the patient to choose the physician's clinical diagnosis. It was calculated as the number of times the patient-selected diagnosis matched the physician diagnosis divided by the number of physician diagnoses. The specificity of the patient-selected diagnosis was defined as how often the patient-selected diagnosis was indeed the physician's clinical diagnosis. Specificity was calculated as the number of times the patient-selected diagnosis matched the physician diagnosis divided by the number of patient-selected diagnoses.

RESULTS

A total of 790 patients began the program. Of these, 424 patients did not complete the entire program, 14 were younger than 18 years, 12 were excluded because there was no progress note supplying a physician's diagnosis, 10 had a diagnosis not capable of being generated by the program, and 2 did not have any clinical diagnosis recorded in the medical record. Therefore, 328 patients completed the program in its entirety and were included in the final analysis. No patient declined an invitation to participate in this study.

There were 163 male and 165 female patients, with a mean age of 48 years (range, 18-76 years). A summary of the data for each individual diagnosis can be seen in Table 2, while the breakdown by age and sex can be seen in Table 3. Table 4 displays the 10 clinical diagnoses that were not capable of being generated by the program during this study; it does not provide a complete list of all diagnoses the program is unable to generate. The most common physician diagnoses were osteoarthritis (n = 157), meniscus tear (n = 131), patellar arthritis (n = 100), patellofemoral-generated pain (n = 83), and anterior cruciate ligament (ACL) tear (n = 29). The program generated a mean 6.6 diagnoses (range, 2-15) per patient, and each patient had a mean 1.7 physician diagnoses (range, 1-4). Patients selected a mean 2 diagnoses (range, 1-9). The program contained the physician's diagnosis 496 times of a total 543 clinical diagnoses, for an overall sensitivity of 91%. There were 496 matches of a total of 2161 diagnoses in the program's differential, for an overall specificity of 23%. The patient-selected diagnosis matched the physician's diagnosis in 315 cases of 543 total clinical diagnoses for a sensitivity of patient-selected diagnosis of 58%. There were 315 matches of a total of 653 patient-selected diagnoses for a specificity of patient-selected diagnosis of 48%.

DISCUSSION

We found that with the aid of an accurate symptom checker, patients were able to correctly identify the cause of their knee pain 58% of the time.

Seventy-two percent of Internet users search for medical information online, and 1 of every 3 adults in the United

TABLE 3
Patient-Selected Diagnoses by Age and Sex

Diagnosis	Mean Age, y	Male, n	Female, n
Anterior cruciate ligament tear	46	22	11
Inflammatory arthritis	49	13	9
Iliotibial band friction syndrome	44	8	8
Lateral collateral ligament tear	41	7	6
Medial collateral ligament tear	43	11	5
Meniscus tear	49	85	68
Osgood-Schlatter	34	5	5
Osteoarthritis	52	40	63
Osteoarthritis exacerbation	52	24	17
Osteochondritis dissecans	0	0	0
Patellar arthritis	48	29	51
Patellar arthritis exacerbation	0	0	0
Patellar chondromalacia/ patellofemoral syndrome	47	16	28
Patellar contusion/saphenous nerve contusion	36	2	6
Patellar instability	35	5	10
Patellar tendinitis	44	6	4
Patellar tendon rupture	42	1	1
Plica	46	14	11
Popliteal cyst	52	4	14
Posterior cruciate ligament tear	34	2	3
Prepatellar bursitis	42	7	8
Quadriceps tendinitis	47	1	5
Quadriceps tendon tear	48	3	2
Stress fracture	57	1	2
Trochlear chondromalacia	50	7	16
Rheumatoid arthritis	52	0	6
Total	48	313	359

TABLE 4
Conditions Not Capable of Being Generated by Program During Study

Ganglion cyst
Traumatic knee strain
Knee effusion
Resolved traumatic bursitis
Gout flare
Quadriceps strain
Postoperative pain
Femoral condylar bruising
Multiple loose bodies
Enchondroma

States has used the Internet to look up information regarding a medical condition that they or someone they know may have.[3] When beginning to research medical information, one can find numerous blogs, forums, and websites filled with testimony from others who share similar symptoms and stories. With the vast amounts of consumer-inputted information, and oftentimes opinions, found on these sites, one must take what is displayed on the screen with caution.

In contrast, there are multiple medical websites and online journals with accurate and up-to-date information available to the consumer. However, as demonstrated by this study, only 58% of patients who completed a symptom checker were able to correctly identify the cause of their knee pain from a list of 2 to 15 diagnoses containing the correct diagnoses 91% of the time, or from a full list of 26 diagnoses, even when the diagnoses were linked to content specifically designed to help them identify the cause of their pain. It would seem that despite the availability of credible medical resources, there is still no guarantee the consumer will interpret this information appropriately when forming their own opinion regarding their medical problem. This highlights the importance of a medical provider performing a physical examination as well as any necessary tests.

The program analyzed in this study differs from other web-based diagnostic programs in that the diagnoses are not generated by simply listing the most frequent diagnoses stratified by sex and age but instead using a process based primarily on the location of the pain as well as whether the mechanism of injury was traumatic or atraumatic. Since we are unaware of any studies testing the accuracy of the alternative method of simply listing the most common diagnoses, we cannot compare the accuracies of the different approaches.

Despite a lower than expected percentage of correct patient-selected diagnoses, patients usually chose diagnoses included within the program's differential and rarely chose a diagnosis that was not included within the differential of the program. There were 15 instances when the patient selected a diagnosis not included within the program's differential. Of these, patients were incorrect 11 times. Of the 4 cases where the patient correctly selected a diagnosis other than those generated by the program, 1 had a known meniscus tear diagnosed by a previous provider and another had an established diagnosis of rheumatoid arthritis. In the other 2 cases, there was no obvious reason for choosing a diagnosis beyond what the program generated. Taken together, this suggests that patients trust symptom checkers and that improved accuracy should be prioritized. However, the accuracy of a program designed to collect a history and generate a diagnosis may not be able to be improved without the ability to gather data from a physical examination or information regarding imaging. Unfortunately, symptom checkers in other specialties have rarely been tested for accuracy. One study showed that WebMD was able to include the correct diagnosis within its differential 70% of the time when attempting to diagnose users with otolaryngological disease.[2]

Interestingly, 38% of patients who had both a clinical and program-generated diagnosis of osteoarthritis did not select it as a cause of their knee pain. The mean age of patients omitting the diagnosis was 65 years (range, 33-76 years). Other studies have demonstrated the patient's hesitance toward accepting a diagnosis of osteoarthritis, especially in younger individuals.[5,8] One explanation may be that osteoarthritis is viewed as a normal part of aging, and younger individuals seek alternative reasons for their knee pain rather than accepting a so-called age-related diagnosis.[6,9,10] As a comparison, the second most common diagnosis (meniscus tear) had agreement between the physician and program 109 times and was omitted by the patient only 17 times (16%).

There are limitations to this study. First, the population was a series of patients presenting to an orthopaedic practice, and the results may not apply a broader group of people searching the Internet to diagnose knee pain. Second, although the diagnoses were made by experienced orthopaedists based on history, physical examination, and radiographs, plus at times advanced imaging, they may not be 100% accurate in every case. When determining the physician diagnosis as the gold standard, studies have shown inconsistent results when attempting to determine clinical accuracy of physician diagnoses for knee pain, citing overall accuracy between 56% and 80%.[7] However, 1 study evaluated the diagnostic accuracy of clinical diagnosis in patients who underwent knee arthroscopy. This study concluded that clinical diagnosis is a reliable method of identifying knee pathology by demonstrating retrospectively that patients presenting to a specialty knee clinic had an overall accuracy, sensitivity, and specificity of clinical diagnosis, respectively, in 99%, 70%, and 99% of patients preoperatively. The cited improvement of results was due to the study being performed in a knee specialty clinic using experienced consultants as the diagnosing clinicians.[7] Another limitation of the study was asking the patient to select all diagnoses they felt could explain their symptoms rather than asking the patient to select the most likely 1 or 2 diagnoses provided by the program. This change would have limited the average number of patient-selected diagnoses during analysis. Additionally, the usefulness of the program is limited by the large differential provided by the program in some of the cases. A large differential improves the sensitivity but not the specificity, where a consistently narrower differential would have likely improved the specificity of the overall program but also would have likely decreased the sensitivity. Furthermore, we were unable to determine whether and how long each patient spent learning about each diagnosis using the linked educational content. Another limitation to the study was excluding those patients who had a diagnosis other than what is capable of being generated by the program. In addition, the program may be better suited for a certain age range or sex; however, that analysis was not performed and therefore cannot be determined. Finally, over half of the ambulatory patients with knee pain did not complete the program in its entirety, and therefore, the results of the study may not be generalizable to the typical patient. It is unknown why these patients did not complete the program in its entirety.

CONCLUSION

We tested the ability of patients to select the cause of their knee pain based on the results of a web-based symptom checker, and found that 58% of 328 patients correctly identified their diagnoses. This program is meant to assist the patient attempting to self-diagnose their knee pain using the Internet as a guide, but our study illustrates the limitations of such a program even when it provides an accurate differential and informative content. Despite these limitations, diagnostic programs may be a valuable resource for those seeking answers to their medical problems.

REFERENCES

1. Bisson LJ, Komm JT, Bernas GA, et al. Accuracy of a computer-based diagnostic program for ambulatory patients with knee pain. *Am J Sports Med*. 2014;42:2371-2376.
2. Farmer SE, Bernardotto M, Singh V. How good is Internet self-diagnosis of ENT symptoms using Boots WebMD symptom checker? *Clin Otolaryngol*. 2011;36:517-518.
3. Fox S, Duggan M. Health online: Pew Internet and American Life Project. 2013. http://pewinternet.org/reports/2013/healthonline.aspx. Accessed April 8, 2015.
4. Fox S, Fallows D. Internet health resources: Pew Internet and American Life Project. 2003. http://www.pewinternet.org/Reports/2003/Internet-Health-Resources.aspx. Accessed February 15, 2015.
5. Gignac MA, Davis AM, Hawker G, et al. ''What do you expect? You're just getting older'': a comparison of perceived osteoarthritis-related and aging-related health experiences in middle- and older-age adults. *Arthritis Rheum*. 2006;55:905-912.
6. Jinks C, Ong BN, Richardson J. A mixed methods study to investigate needs assessment for knee pain and disability: population and individual perspectives. *BMC Musculoskelet Disord*. 2007;8:59.
7. Nickinson R, Darrah C, Donell S. Accuracy of clinical diagnosis in patients undergoing knee arthroscopy. *Int Orthop*. 2010;34:39-44.
8. Paskins Z, Sanders T, Hassell AB. Comparison of patient experiences of the osteoarthritis consultation with GP attitudes and beliefs to OA: a narrative review. *BMC Fam Pract*. 2014;15:46.
9. Paskins Z, Sanders T, Hassell AB. What influences patients with osteoarthritis to consult their GP about their symptoms? A narrative review. *BMC Fam Pract*. 2013;14:195.
10. Sanders C, Donovan J, Dieppe P. The significance and consequences of having painful and disabled joints in older age: co-existing accounts of normal and disrupted biographies. *Sociol Health Illness*. 2002;24:227-253.