Approximately 43% per 100,000 people are diagnosed in the latter stages of cancer. There are multiple reasons for the high rate of late-stage diagnosis, such as missing patient history. Often the patient is lacking in knowledge, causing a situation where they misinterpret symptoms or even notice at times. From the doctor’s side, certain symptoms may be difficult to interpret, and it may be difficult to find tumors using medical images. There are certain types of cancers whose symptoms are similar to other less lethal diseases. Late diagnoses are commonplace as well, with some cancers, such as colorectal and cervical cancers (approximately 50% of these are diagnosed late) and breast cancer (approximately a third of these are diagnosed in the latter stages) having very high rate of late diagnosis. The fastest procedure for identifying cancer is through a medical image, but due to the difficulties that doctors may have in identifying the tumor, further studies are needed to confirm the tumor. In turn, machine learning algorithms can help identify malignant tumors via images, even highlighting the area with the possible tumor itself for early diagnosis.

Early diagnosis is crucial to the survivability of cancer patients. Not only does early diagnosis help patients save on costs (estimated to be between 8-11% depending on the stage) by allowing cheaper and effective procedures, but it also increases the survival rate of the patients. Approximately 90% of women who are diagnosed in the earlier stages of breast cancer live more than five years after the tumor is removed compared to the 15% of women who are diagnosed at the most advanced stage of cancer. Similarly, 70% of lung cancer patients survive for more than one year if diagnosed at the earliest stage in comparison to the 14% of patients diagnosed with at the most advanced stage. Considering the potential savings together with the increased survivability from early diagnosis, a cost between $500 and $1,000 is more than appealing when including the fact that cancer can possibly place a patient into bankruptcy due to the expensive medicine they may need to buy while at the latter stages of cancer. By combining early screenings together with machine learning algorithms, one will not only be able to capture tumors at their earlier stages, thus saving time, money, and increasing survivability, but also prevent the pain caused by cancer at its later stages.

Machine learning is a branch of artificial intelligence (AI) that focuses on using data to imitate how humans learn and make predictions with higher speed and accuracy. Neural Networks, which are a set of algorithms that focuses on imitating the way in which the neurons within the body function to learn and a branch underneath machine learning, function to make very fast prediction and process high volume of data to increase accuracy. To process the images of cancer patients, a Convolutional Neural Network (CNN) will be required. The key to this solution is not based on the architecture of the machine learning algorithm, but rather the data available. With images containing labels regarding the malignancy of the tumor, the stage, the best possible treatment, one could train a Convolutional Neural Network that efficiently predicts what may be the best next procedure for treating the tumor of the patient. This allows one to gather data retroactively (noting down the most effective procedures, collecting notes on other factors that may have had an impact on the cancer, etc.) and use that data to train the model to input an image and output a prediction of whether the tumor is malignant or benign, the stage of the tumor, and the next best possible procedure to take.

This proposed data product solution will classify and provide information regarding tumors based on image while independent of the imaging modality. Although CT-Scans are the most common form of image cancer screening, there are other imaging modalities such as MRI which can potentially be used to identify cancer. This data product will save time for doctors who may have to closely examine the image or extract a sample of tissue to observe whether the sample tissue contains cancer cells or not. For the patient, it will be less stressful waiting for the doctor’s diagnosis while removing possible human error which is the main cause of late-stage diagnosis (from both the patient’s and the doctor’s side). From the healthcare industry perspective, this may provide some confidence in the required prior authorization for further cancer tests. Although this algorithm may be a good case for doctors to follow through on further procedures, minimizing the number of false negatives while also providing further evidence for further testing towards the insurance company, there may be some doubt in regard to trusting a machine to predict cancer.

Research shows that there have been machine learning algorithms with the ability to even identify and predict a woman’s potential breast tumor four years before it developed. Despite this, there are little to no alternatives outside of special products or proof of concepts. The main factor to blame is that until recently, hospitals and even pharmaceutical companies have gathered their data within data silos that make it difficult to process the data. Although there are the imaging data standards, such as the Picture Archiving and Communication System (PACS) and the Digital Imaging and Communications in Medicine (DICOM), they are rarely processed to contain accurate labels, annotations, and other factors for developing machine learning models. This has caused for large technology companies to unite with hospital or medical research institutions to provide the three products each at different stages. Google’s DeepMind has a research partnership with the radiotherapy department, IBM’s Watson AI, and finally the Digital Mammography DREAM challenge.

Google’s DeepMind model already has the capability to perform better than doctors at breast cancer screening. It does this by using both 2D and 3D mammography images from approximately 30,000 women. As opposed to the proposed data product, the DeepMind model focuses on a singular organ/body part which limits the model to screening only a specific type of cancer. Although Watson has been proposed as a solution for assisting doctors in determining the best next procedure for cancer patients, it has had difficulty in learning different types of cancer. Finally, the Digital Mammography DREAM challenge was simply a challenge that focused on breast cancer patients and novel algorithms to assist breast cancer patients. This challenge was purely research based. From these products, specially the third, one can obtain samples of possible machine learning algorithms. The intended technologies for this data product is the Imaging Data Common’s Google BigQuery database, python’s pydicom and TensorFlow libraries.

As a Christian, my perspective to help other through any knowledge or tool possible is always a priority. We are called to love others in the same way that we love ourselves. Therefore, this brings more motivation to develop a product that not only helps doctors by freeing up some of their time, but by also increasing the rate of early cancer diagnosis. The project would not change by my Christian world view because my Christian worldview has been the motivation for this project.

Source(s):

*12 Startups Bringing Artificial Intelligence To Cancer Diagnostics And Therapeutics*. (2016, September 15). CB Insights Research. https://www.cbinsights.com/research/ai-startups-fighting-cancer/

Bazargan, M., Lucas-Wright, A., Jones, L., Vargas, R., Vadgama, J. V., Evers-Manly, S., & Maxwell, A. E. (2015). Understanding Perceived Benefit of Early Cancer Detection: Community-Partnered Research with African American Women in South Los Angeles. *Journal of Women’s Health*, *24*(9), 755–761. https://doi.org/10.1089/jwh.2014.5049

Bi, W. L., Hosny, A., Schabath, M. B., Giger, M. L., Birkbak, N. J., Mehrtash, A., Allison, T., Arnaout, O., Abbosh, C., Dunn, I. F., Mak, R. H., Tamimi, R. M., Tempany, C. M., Swanton, C., Hoffmann, U., Schwartz, L. H., Gillies, R. J., Huang, R. Y., & Aerts, H. J. W. L. (2019). Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA: A Cancer Journal for Clinicians*. https://doi.org/10.3322/caac.21552

Crosby, D., Bhatia, S., Brindle, K. M., Coussens, L. M., Dive, C., Emberton, M., Esener, S., Fitzgerald, R. C., Gambhir, S. S., Kuhn, P., Rebbeck, T. R., & Balasubramanian, S. (2022). Early detection of cancer. *Science*, *375*(6586). https://doi.org/10.1126/science.aay9040

*Digital Mammography DREAM Challenge*. (n.d.). Sage Bionetworks. Retrieved September 8, 2022, from https://sagebionetworks.org/research-projects/digital-mammography-dream-challenge/

Hale, C. (2020, January 2). *An instant 2nd opinion: Google’s DeepMind AI bests doctors at breast cancer screening*. Fierce Biotech. https://www.fiercebiotech.com/medtech/instant-second-opinion-google-s-deepmind-ai-bests-doctors-at-breast-cancer-screening#:~:text=Google%27s%20DeepMind%20team%20showed%20that

IBM Cloud Education. (2020a, July 15). *What is Machine Learning?* IBM. https://www.ibm.com/cloud/learn/machine-learning

IBM Cloud Education. (2020b, August 17). *What are Neural Networks?* Www.ibm.com; IBM. https://www.ibm.com/cloud/learn/neural-networks

*Individual Heath Insurance Options*. (n.d.). Www.cancer.org. https://www.cancer.org/treatment/finding-and-paying-for-treatment/understanding-health-insurance/private-insurance-options.html

Kakushadze, Z., Raghubanshi, R., & Yu, W. (2017). Estimating Cost Savings from Early Cancer Diagnosis. *Data*, *2*(3), 30. https://doi.org/10.3390/data2030030

Kaur, S., Baine, M. J., Jain, M., Sasson, A. R., & Batra, S. K. (2012). Early diagnosis of pancreatic cancer: challenges and new developments. *Biomarkers in Medicine*, *6*(5), 597–612. https://doi.org/10.2217/bmm.12.69

Lee, S. (2014). *Benefits and limitations of regular cancer screening*. Canadian Cancer Society. https://cancer.ca/en/cancer-information/find-cancer-early/screening-for-cancer/benefits-and-limitations-of-regular-cancer-screening#:~:text=Cancer%20screening%20helps%20find%20cancer%20before%20it%20spreads%20when%20it

Ross, C., & Swetlitz, I. (2017, September 5). *IBM pitched its Watson supercomputer as a revolution in cancer care. It’s nowhere close*. STAT; STAT. https://www.statnews.com/2017/09/05/watson-ibm-cancer/

Savage, N. (2020). How AI is improving cancer diagnostics. *Nature*, *579*, S14–S16. https://doi.org/10.1038/d41586-020-00847-2

Smith, & Smith. (2022, May 11). *8 Reasons Why Your Doctor Missed a Cancer Diagnosis*. STL.News. https://www.stl.news/8-reasons-why-your-doctor-missed-a-cancer-diagnosis/519494/

Staff, N. (2022, March 22). *Can Artificial Intelligence Help See Cancer in New Ways? - National Cancer Institute*. Www.cancer.gov. https://www.cancer.gov/news-events/cancer-currents-blog/2022/artificial-intelligence-cancer-imaging

*Stage at Diagnosis | Cancer Trends Progress Report*. (2019). Cancer.gov. https://progressreport.cancer.gov/diagnosis/stage

The Lancet. (2010). Late-stage cancer detection in the USA is costing lives. *The Lancet*, *376*(9756), 1873. https://doi.org/10.1016/s0140-6736(10)62195-2

*Ultrasound for Cancer*. (n.d.). Www.cancer.org. https://www.cancer.org/treatment/understanding-your-diagnosis/tests/ultrasound-for-cancer.html#:~:text=Ultrasound%20cannot%20tell%20whether%20a

Virnig, B. A., Baxter, N. N., Habermann, E. B., Feldman, R. D., & Bradley, C. J. (2009). A Matter Of Race: Early-Versus Late-Stage Cancer Diagnosis. *Health Affairs*, *28*(1), 160–168. https://doi.org/10.1377/hlthaff.28.1.160

Xie, Y., Bagby, T. R., Cohen, M., & Forrest, M. L. (2009). Drug delivery to the lymphatic system: importance in future cancer diagnosis and therapies. *Expert Opinion on Drug Delivery*, *6*(8), 785–792. https://doi.org/10.1517/17425240903085128