Dataset Link: https://www.kaggle.com/datasets/falgunipatel19/ text-publication-classification

For Biomedical text document classification, abstract and full papers(whose length less than or equal to 6 pages) available and used. This dataset focused on long research paper whose page size more than 6 pages. Dataset includes cancer documents to be classified into 3 categories like 'ThyroidCancer','ColonCancer','Lung_Cancer'. Total publications=7569. it has 3 class labels in dataset. number of samples in each categories: colon cancer=2579, lung cancer=2180, thyroid cancer=2810

In [1]:
```
#Import necassary libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.feature_extraction.text import TfidfVectorizer
from nltk.corpus import stopwords
from nltk.util import ngrams
from nltk.stem import WordNetLemmatizer
from nltk.sentiment.vader import SentimentIntensityAnalyzer
import re
from wordcloud import WordCloud
from textblob import TextBlob
from nltk import word_tokenize,sent_tokenize
import spacy
nlp = spacy.load('en_core_web_sm', disable=['ner'])
import string
from nltk.corpus import stopwords
from nltk.stem.porter import PorterStemmer
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
import warnings
warnings.filterwarnings('ignore')
```

```
2022-09-10 21:09:14.510426: I tensorflow/core/util/util.cc:169] oneDNN custom
operations are on. You may see slightly different numerical results due to flo
ating-point round-off errors from different computation orders. To turn them o
ff, set the environment variable `TF_ENABLE_ONEDNN_OPTS=0`.
2022-09-10 21:09:14.550292: W tensorflow/stream_executor/platform/default/dso_
loader.cc:64] Could not load dynamic library 'libcudart.so.11.0'; dlerror: lib
cudart.so.11.0: cannot open shared object file: No such file or directory
2022-09-10 21:09:14.550324: I tensorflow/stream_executor/cuda/cudart_stub.cc:2
9] Ignore above cudart dlerror if you do not have a GPU set up on your machin
e.
2022-09-10 21:09:16.405241: W tensorflow/stream_executor/platform/default/dso_
loader.cc:64] Could not load dynamic library 'libcuda.so.1'; dlerror: libcuda.
so.1: cannot open shared object file: No such file or directory
2022-09-10 21:09:16.405292: W tensorflow/stream_executor/cuda/cuda_driver.cc:2
69] failed call to cuInit: UNKNOWN ERROR (303)
2022-09-10 21:09:16.405322: I tensorflow/stream_executor/cuda/cuda_diagnostic
s.cc:156] kernel driver does not appear to be running on this host (vinod-Vost
ro-3400): /proc/driver/nvidia/version does not exist
```

In [2]:
```
#Read the data using the pandas
data=pd.read_csv('/home/vinod/Downloads/alldata_1_for_kaggle.csv',encoding='la
data.head()
```

Out[2]:

	Unnamed: 0	0	a
0	0	Thyroid_Cancer	Thyroid surgery in children in a single insti...
1	1	Thyroid_Cancer	" The adopted strategy was the same as that us...
2	2	Thyroid_Cancer	coronary arterybypass grafting thrombosis ï¬^b...
3	3	Thyroid_Cancer	Solitary plasmacytoma SP of the skull is an u...
4	4	Thyroid_Cancer	This study aimed to investigate serum matrix ...

In [3]:
```
#data information in the dataset
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7570 entries, 0 to 7569
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   Unnamed: 0  7570 non-null   int64
 1   0           7570 non-null   object
 2   a           7570 non-null   object
dtypes: int64(1), object(2)
memory usage: 177.5+ KB
```

In [4]:
```
#Check the data shape of the dataset
data.shape
```

Out[4]:
```
(7570, 3)
```

In [5]:
```
#Check the null values in the dataset
data.isna().sum()
```

Out[5]:
```
Unnamed: 0    0
0             0
a             0
dtype: int64
```

Data Preprocessing

In [6]:
```
#Let's remove the unnecassary columns in the dataset
data.drop(['Unnamed: 0'],axis=1,inplace=True)
data.head().style.background_gradient(cmap='winter')
```

Out[6]:

	0	0
0	Thyroid_Cancer	Thyroid surgery in children in a single institution from Osama Ibrahim Almosallama Ali A Alsobhib Saud AlShanafeybFrom the aDepartment of Surgery College of Medicine Qass bDepartment of Surgery King Faisal Specialist Hospital and Research Center Riyadh Sau Specialist Hospital and Research Center Riyadh Saudi Arabia Correspondence Dr Osa College of Medicine Qassim University PO Box Buraidah Al Qassim Saudi Arabia osama_ia Citation Almosallam OI Aseeri A Alhumaid A AlZahrani AS Alsobhi S AlShanafey S Thyro Ann Saudi Med Received January Accepted May Published August Copyright Copyright Ar access under the Creative Commons AttributionNonCommercialNoDerivatives International be accessed at httpcreativecommons licensesbyncnd40Funding NoneBACKG scarceOBJECTIVE Analyze outcome data on thyroid surgery in a pediatric populationDI health care institutionPATIENTS AND METHODS We collected demographic and clinical dat surgery in the period to Descriptive data are presentedMAIN OUTCOME MEASURES complications length of stay and radioactive iodine treatment and recurrencesSAMP thyroidectomy procedures were females and the mean age at operation was years and wer type There was no history of radiation exposure Eightyone patients had fine needl histopathology in of cases Sixtysix patients had malignant cancer papillary of patients who h and had distant metastases to the lung Procedures included total thyroidectomy hemithyroid Twentythree patients developed hypocalcemia permanent and had unilateral recurrent followed up for a mean duration of months median months Of patients with thyroid cancer Malignancy is the commonest indication for thyroid surgery in children and FNA is highly di nerve injury are significant complications The recurrence rate in thyroid cancer is LIMITAT Noneoriginal ANN SAUDI MED JULYAUGUST WWWANNSAUDIMEDNET 0cThyroid disease children compared to adults The prevalence of palpable thyroid nodules in children ranges fr is the most common endocrine malignancy in children accounting for of pediatric cancers in in adolescents aged year2 The most common indication for thyroid surgery in children var for malignant conditions is rising38 Data in children throughout the world are relatively scarc clinical data and outcome of thyroid surgery in a large series of children treated at a sing Research Center KFSHRC in RiyadhPATIENT AND METHODS With the approval of the medical records of all patients years old and younger who underwent a thyroid surgery l elected to include patients up to the year to ensure a reasonable followup period Patients operating room log for all procedures involving the thyroid gland for the specified age groupI outcomes were collected Specific data that were obtained included age at operation gende radiation exposure presence of multiple endocrine neoplasia type MEN thyroid functi ultrasound presence of lymph nodes metastasis or distant metastasis fine needle a histopathology and length of followup Outcomes analyzed were postoperative complications transient or permanent recurrent laryngeal nerve paralysis wound infection and hematoma and recurrences Thyroid procedures in this series included hemithyroidectomy subtotal tota performed by either an endocrine adult surgeon or a pediatric surgeon No intraoperative procedures were performed by adult endocrine surgeons but lately a combined approach w endocrine surgeons collaborated in such cases proceduresthe normal range in ou hypocalcemia was identified if it lasted for less than months while permanent hypocalc remained below normal range and the patient continued on calcium supplementation after history of MEN underwent genetic testing of the RET protooncogene to confirm the di thyroidectomy had a preoperative and postoperative vocal cords assessment at the Otolary and comparisons were conducted using the t test for continuous proportionsRESULTSBetween and patients underwent surgical procedures patients unde institution Eighty patients were females The mean age at operation was years median yea thyroidectomy was thyroid nodule which was present in of cases Table The mean SD size associated with MEN syndromes The final pathology in two patients with MEN syndrome s remaining patients had prophylactic procedures before development of MTC None of Eightyone patients FNA which correlated with the final histopathology in of cases There were Graves disease which did not require FNA The remaining cases underwent FNA at anotl institution or they came for completion thyroidectomy with documented pathology for malign hospitalThe most common diagnoses included papillary thyroid cancer and multinodular goit in patients IndicationNodulen MEN prophylaxisHyperthyroidismMultinodular goiterCompletic calcium levels below Data are number original PEDIATRIC THYRO WWWANNSAUDIMEDNET 0cnodule Table Surgical procedures included total thyroidectom and subtotal thyroidectomy Neck dissection was performed in patients Operative cor common complication was hypocalcemia transient permanent and Table Thyroid pathology i tissueColloid noduleCystAdenomaThyroiditisGraves diseaseThyroid cancerPapillaryFol number Table Benign and malignant lesions in patientsBenignn37Malignantn66 P value A noduleHypocalcemiaRecurrent laryngeal nerve palsyBleedinghematomaWound infectionTrac stay daysMEN recurrent laryngeal nerve palsy transient permanent all were unilateral Tab node metastasis and patients had distant metastases to the lung None of the patients devel

0

tracheal injury Patients were followed up for a mean of months median range months radioa
with malignant lesions patients had recurrences were local recurrences and were local *a*
received radioactive iodine RAI before and after recurrence One case was low risk bef
recurrence One case had medullary thyroid cancer so did not receive RAI In the remaining fiv
patients received RAI before or only after a recurrence All local recurrences underwent rese
up There was no mortality in this study DISCUSSIONThe most common indication for thyroic
correlates with previously published reports in the pediatric population35 Children with thyroid
of developing thyroid cancer compared to adults910 The high incidence of malignancy in this
should be carefully evaluatedFNA is a valuablemethod for preoperative evaluation of thyro
routine use of FNA in children including the need for sedation sampling errors and the limite
Many previous studies reported high sensitivity and specificity of FNA in evaluating thyroid
findingsOur data showed lymph node metastasis in of thyroid cancer cases which supp
frequently present with more extensive disease than adults Lymphnode involvement at diag
of adults with differentiated thyroid cancer1523 Because our hospital is the largest referra
cases this may explain the large number of lymph node and distant metastasis In this cohor
thyroidectomy in children is hypoparathyroidism with an incidence ranging between to which
SAUDI MED JULYAUGUST WWWANNSAUDIMEDNET 0ccorresponds with our results of wh

study found that total thyroidectomy central and bilateral neck dissection Graves disease and

after thyroid surgery3 In this cohort postoperative hypocalcemia was noted more in
significance Moreover there was no significant difference between benign and malignant
recurrent laryngeal nerve injury or overall complications a finding that was reported previousl
an inverse relationship between surgeon volume and complication rates2728 but similar data
found that highvolume endocrine surgeons have better outcomes and shorter lengths
parathyroidectomy in children compared to pediatric surgeons general surgeons or otolar
concluded that a collaborative approach between pediatric and endocrine surgeons would h
to suggest that a combined approach with endocrine and pediatric surgeons in addition to p
of children with surgical thyroid disease given the low number of pediatric patients4 Ou
approaches given the late adoption of the combined approach The recurrence rate for thy
varied widely in reported studies ranging from to while it was in this cohort Only a few studi
node involvement multiple nodules male gender younger age histologic subtype and advance
recurrence17233033 In this study of patients with malignant lesions received RAI A
indications of postoperative RAI treatment in lowrisk patients the current recommendatio
RAI3436There are some limitations to this study The retrospective nature may affect the valid
cases in some categories did not enable us to compare groups and explore predictors relativ
adds to the scarce data on thyroid surgery in pediatric age group Malignancy is the commone
FNA is highly diagnostic Hypocalcemia and recurrent laryngeal nerve injury are significant
rare but recurrence is not uncommon and a significant number of patients with malignant ca
THYROID SURGERYANN SAUDI MED JULYAUGUST WWWANNSAUDIMEDNET
McLaren GD Nichaman MZ Iodine and goiter in children Pediatrics Ries LAG Melbert D Krap

et al SEER Cancer Statistics Review Bethesda National Cancer Institute Based on Novemb

PT Gaz RD Hodin RA Parangi S Randolph GW et al Pediatric thyroidectomy in a high
postoperative hypocalcemia J Pediatr Surg Aug5081316 Wood JH Partrick DA Barham HP B
thyroidectomy a collaborative surgical approach J Pediatr Surg May4658238 Scholz S Smi

Thyroid surgery at Childrens Hospital Boston a 35year singleinstitution experience J Pediat

Thyroid nodules and cancers in children Pediatr Endocrinol Rev Sep611423 Hameed F
adolescent thyroid cancer J Paediatr Child Health LugoVicente H Ortiz VN Irizarry
management in the era of fine needle aspirationJ Pediatr Surg Mussa A De Andrea M Motta
Malignancy in Children with Thyroid Nodules J Pediatr Oct167488692 Amirazodi E Propst E
thyroid FNA biopsy Outcomes and impact on management over years at a tertiary care ce
Cramer HM Chen S Wu HH Histologic and clinical followup of thyroid fineneedle aspirates in
Decoppi P Pierro A Brain C Hindmarsh P Butler G et al Thyroid Surgery in Children Clin
Kundel A Thompson GB Richards ML Qiu LX Cai Y Schwenk FW et al Pediatric Endocrine S

J Clin Endocrinol Metab February Jiang W Newbury RO Newfield RS Pediatric thyroid

institutional experience features and over a 10year period Int J Pediatr Endocrinol Burke JF S

Surgery at a Tertiary Medical Center Surg Res AlQahtani KH Tunio MA Al Asiri M Aljohan

treatment outcomes of differentiated thyroid cancer in Saudi children and adults J Otolaryngo
DJ Verrijn Stuart AA Lodewijk L Valk GD Van der Zee DC et al Postoperative Complica
Young Patients With Multiple Endocrine Neoplasia Type Medicine Baltimore 20159429e11

Angelos P Reynolds M Total thyroidectomy for benign disease in the pediatric patientfeasi

PH Ko CY Yeh MW Surgeon volume as a predictor of outcomes in inpatient and outpatien

HM Tielsch JM Powe NR Gordon TA Udelsman R The importance of surgeon expe

		0
		thyroidectomy Ann Surg Tuggle CT Roman SA Wang TS Boudourakis L Thomas D Udels
		operating on our children Surgery Dec144686977 Park S Jeong JS Ryu HR Lee C Park JH K Children and Adolescents27Year Experience in the Yonsei University Health System J Ko Kollars JP Moir CR Papillary thyroid carcinoma in children risk factors and complication Sugino K Mimura T Nagahama M Kitagawa W Shibuya H et al Pediatric differentiated thy disease free survival BMC Cancer D Danese Gardini A Farsetti A Sciacchitano S Andreoli M I adolescents Eur J Pediatr Astl J Chovanec M Lukes P Katra R Dvorakova M Vlcek P
		adolescents years experience surgery of pediatric thyroid lymph node metastases carcin
		Rangarajan V Nair N Nadkarni MS Pai PS Dcruz AK et al Pediatric thyroid cancer J Su Matovic M Milovanovic Z et al Surgical management of welldifferentiated thyroid carcinoma i of a single institution in Serbia Endocr J Scheumann GF Gimm O Wegener G Hundeshagen management of locoregional in papillary thyroid cancer World J Surg Shi RL Qu N Yang SW
		lymph node metastasis using a differentiated thyroid cancer risk model Onco Targets The
		Ryan JJ Grant CS et al Papillary thyroid carcinoma in children and adults longterm fo institution during three decades Surgery Collini P Mattavelli F Pellegrinelli A Barisella M Fe thyroid gland of childhood and adolescence Morphologic subtypes biologic behavior and cases treated at a single institution during a 30year period Am J Surg Pathol BorsonChazot JL Predictive factors for recurrence from a series of children and adolescents with differen HD Bauer AJ Isaza A MostoufiMoab S Kazahaya K Adzick NS Surgical management of pe
		thyroidectomy at the Childrens Hospital of Philadelphia highvolume Pediatric Thyroid Center J
		De Coppi P Thyroidectomy in Children InPediatric Surgery pp Springer Berlin Heidelberg F Benvenga S et al Management Guidelines for Children with Thyroid Nodules and Differ Association Guidelines Task Force on Pediatric Thyroid Cancer THYROID Volume Number SAUDI MED
1	Thyroid_Cancer	" The adopted strategy was the same as that used in prior years [] and is based on four e subsets The first query QPub_plain is based on a plaintext search in PubMed titles and s usi relies on the PubMed indexing scheme using MeSH terms and results are m
		QWoS_restricted is based on a plaintext search in WoS restricted to the two research areas
		Services The fourth query QWoS_filtered is based on the same plaintext search used in Wo
		Archeology Dance Zoology etc and the two research areas of the previous query It nonPubMedindexed papers that are supposed to be caught by the two PubMed quer citations was performed by the two section editors to select candidate best papers Followi best papers were then individually reviewed and rated by both section editors the chief edit
		reviewers from the international Medical Informatics community Based on the reviewe
		committee then selected the best papers of the year in the decision support domainIMI Thieme Verlag KG 0cReview Results The literature search has been performed on Jan distributed as follows for QPub_plain for QPub_indexed for QWoS_restricted and for QWo PubMed and from WoS Compared to the previous year the global query retrieve independently performed by both section editors based on the title and of papers not rejecte two editors to achieve a final selection of candidate best papers After the external review o
		three of them as best papers for [] Table They are discussed in the next section and
		AppendixDiscussion and OutlookIn the first paper Hendriks [] propose an approach to th certainly builds on already existing approaches but which is systematically conducted in orde guidelines They promote the formalism of clinical decision trees CDTs as they are both clir and computerinterpretable thus suitable for implementation in datadriven CDSSs The disam by the formal unequivocal specification of data items used as decision criteria using internat and second by the representation of guideline knowledge as CDTs The method is applied to were built involving a total of data items among which could not be linked to standard te certain criteria which could be subjective or had multiple definitions The resulting knowledg application where it can be interactively browsed or automatically executed By modeling guid
		in the sharing of encoded knowledgeIn the second paper KamiÅ¡aliÄ› [] tackled the i processes used for managing chronic diseases and their execution in CDSSs They a therapeutic management of chronic diseases like those known to increase the cardiovascu strategy dosage adaptation and intolerance management To handle these different aspec extended Timed Transition Diagram eTTD With eTTDs they illustrate the multilevel ar contents of arterial hypertension management guidelines This detailed demonstration management can be formalized to develop a CDSS could certainly be used in other medical conceptual and practical framework to help assess confidence in predictive tools GRASP f method to look for evidence from the published literature and an analysis grid It standar

0
associated to a predictive tool and the grading of its level of proof Three phases of evaluatio the tool to assess both its internal and external validity ii during the implementation to assess implementation to assess its effectiveness and safety In each phase the level of evid qualitative summarizes the direction of evidence positive negative mixed This grid can be co the CONSORT statement for clinical trials However it gives a rigorous methodology for a c extended to all kind of CDSSs It might be a useful tool to extend the evidencebased cultu three best papers selected for the Decision Support section of the edition of the IMIA Y literature review deserve to be cited Some of them deal with the personalization of decis approach to develop personalized care plans that comply with clinical practice guidelines situations Jafarpour [] propose a solution to dynamically manage the conflicts that can rise introduce the use of health information technology involving multiple criteria decision to sup Interestingly other works promote the creation and sharing of operational knowledge base transform the textual STOPPSTART criteria into unambiguous definitions mapped to n EUCAST expert rules as an ontology and production rules to detect antimicrobial therapies a knowledge base that can compete with commercial ones Replacing humans is another topic c virtualize a doctor the automatic acquisition of data through sensors and speech recognit Rozenblum et al[] propose a machine learning method to generate clinically valid alerts CDSS is another key point Kannan [] propose a method for a CDSS design to best meet a pr Design alerts may also avoid rejection of CDSSs by caregivers Fernandes [] created alg notifications delivered to healthcare professionals Amrose et al [] tried to understand in real actions they triggered Finally it is always interesting to obtain varied evaluation resul evaluated Watson for Oncology in thyroid carcinoma and reported a concordance rate with
the tool As evidenced by the number and the variety of works around decision suppo
selection highlighted pragmatic works that promote the transparency and sharing of the IMI 0cTable Best paper selection of s for the IMIA Yearbook of Medical Informatics in th
alphabetical order of the first authors surname Section Decision Support\uf0a7 Hendriks MP
MJC Strobbe LJA Merkus JWS Zonderland HM Smorenburg CH Jager A Siesling S Transfor Into DataDriven Clinical D
KamiÅ¡aliÄ❯\tA\tRia±o\tD\tKert\tS\tWelzer\tT\tNemec\tZlatolas\tL\tMultilevel\tmedical\tknowled
for chronic diseases Data Knowledge Engineering \uf0a7 Khalifa M Magrabi F Gallego
grading and assessment of predictive tools for clinical decision support BMC Med Inform support tools as well as the grading of their utility The ultimate goa themAcknowledgementWe would like to thank all the present and past editorial boards of th Adrien Ugon for their support as well as the reviewers for their participation to the selec section We cannot end this synopsis without a meaningful thought for our colleague and frien to tackle the tasks of a Decision Support section coeditor but passed away in last Decembe Jankovic I Chen JH Clinical Decision Support and Implications for the Clinician Burnou Contributions on Clinical Decision Support from the Literature Yearb Med Inform Aug28113 van der Sangen MJC Strobbe LJA Merkus JWS Transformation of the National Breast Can
Trees JCO Clin Cancer Inform KamiÅ¡aliÄ❯ A Ria±o D Kert S Welzer T Nemec Zlatolas
support medical practice for chronic diseases Data Knowledge Engineering Khalifa M M
evidencebased grading and assessment of predictive tools for clinical decision support B Sarigul B Arvanitis TN Lindman P Chen R A Collaborative Platform for Management of Chro
Care Plans Comput Struct Biotechnol J Jafarpour B Raza Abidi S Van Woensel W Raz
practice guidelines to provide decision support for comorbid conditions Artif Intell Med Ben S PARS a system combining semantic technologies with multiple criteria decision aiding for sup Huibers CJA Sallevelt BTGM de Groot DA Boer MJ van Campen JPCM Davids CJ C algorithms for software implementation A multidisciplinary consensus procedure Int J Med Campos M Palacios F Impact of expert knowledge on the detection of patients at risk of support systems J Biomed Inform M¼ller L Gangadharaiah R Klein SC Perry J Bernstein G for community driven diagnostic decision support system development BMC Med Inform Canbay A Menrad K Heider D The virtual doctor An interactive clinicaldecisionsupport prediction of diabetes Artif Intell Med Rozenblum R RodriguezMonguio R Volk LA Fors Learning System to Identify and Prevent Medication Prescribing Errors A Clinical and Cost A Kannan V Basit MA Bajaj P Carrington AR Donahue IB Flahaven EL User stories as lig support development J Am Med Inform Assoc Fernandes CO Miles S Lucena CJP Cowan with Alarm Fatigue in Hospital Environments Because of Sensory Overload Algorithm D 20192111e15406 Amroze A Field TS Fouayzi H Sundaresan D Burns L Garber L et al Use Logs to Identify Physician Actions Following Noninterruptive Alert ing Descriptive Study J Kim JM Kim EH Kim K Pak K Concordance in postsurgical radioactive iodine therapy recom clinical practice in patients with differentiated thyroid carcinoma Cancer Correspondence t L©onard de Vinci rue Marcel Cachin Bobigny FranceEmail catherineduclosaphpfr IMIA Y Considerations on Clinical Decision Support from the Literature 0cAppendix Content Sum

		0
		Section of the IMIA YearbookHendriks MP Verbeek XAAM van Vegchel T van der Sangen I Smorenburg CH Jager A Siesling STransformation of the National Breast Cancer Guideline Cancer Inform May3114Since clinical practice guidelines are still narrative and described in la to model complex guidelines as datadriven clinical decision trees CDTs that could be still hum implementation in decision support systems The Dutch national breast cancer guidelin characterize the patient and the tumor and represent decisional criteria were encoded u coding systems related to breast cancer when feasible In total CDTs were necessary to cov all data items could be coded using existing classification and coding systems All CDTs repre guidelines could be transformed as systematically constructed modular datadriven CDTs tha decision support applicationKamiÅ¡aliÄ› A Ria±o D Kert S Welzer T Nemec Zlatolas LMultile medical practice for chronic diseasesData Knowledge Engineering This research is focus medical processes involved in chronic diseases management which can be viewed as a proc An intuitive easy and effective mechanism for medical knowledge formalization is propos Transition Diagram eTTD This formalism allows for the consistent representation of three taken into account in the prescription and adaptation of longterm treatment therapy strategy be manually applied to build eTTDs from clinical practice guidelines eTTDs implementatio guidelines for the therapeutic management of arterial hypertension The obtained mode development of decision support systems involving medical proceduresKhalifa M M evidencebased grading and assessment of predictive tools for clinical decision supportBMC a clinical predictive tool in clinical practice should be guided by its correctly assessed effectiv a conceptual and practical framework to Grade and Assess Predictive tools GR evidencebased system to support their search for and selection of efficient predictive tools based on published evidence across three dimensions phase of evaluation level of evidence tool is based on the phase of evaluation that gets the hightest grade supported by the supports a positive This framework was successfully applied to five predictive tools GRAS data base that documents the evidence of predictive tools IMIA Y
2	Thyroid_Cancer	coronary arterybypass grafting thrombosis ï¬ˆbrin ï¬ˆbrinogen mutationIntroduction Intraoper harvesting is very rareCase Report We present a case of a 60yearold male patient with multiv nonST elevation acute coronary syndromeand type2 diabetes mellitus whombilateralintraoperative SV thrombosis occurred during graft harvesting Routinethromb cancer was excluded Compared with healthy controls we observed prolonged ï¬ˆbrin c reï¬(ected by endogenous thrombin potential Scanning electronmicroscopy of the thrombos layer on theclot surface with a solid mass of unusually compressed platelets and eryt ï¬ˆbrinogen and factor F XIII polymorphismsand was found to be heterozygous for ï¬ˆbr TConclusion ï¬ˆbrinogen HaeIII and FXIII Val34Leu polymorphisms are reï¬(ected inreduced might contribute to intraoperative SV thrombosis during vascular grafting procedures Car failure after bypass proceduresIntroductionCoronary artery bypass grafting CABG is a metho multivessel disease anddiabetes mellitus DM Although arterial grafts are preferredin se leftinternal thoracic artery LITA to bypass the left anteriordescending artery LAD and to p often chosen vascular graft the greatsaphenous vein SV offers decent durability and is easy t of caseswithin the ï¬ˆrst months and as many as may occludewithin ï¬ˆrst to weeks environment with disruption of bloodï¬(ow in vasa vasorum damage to the adventitia hypoxi focal endothelialdisruption2 Acute SV graft failure is usually a result of graftthron failuregrafttarget vessel disproportion etc may be caused byhypercoagulabilityreceivedMarc Thieme Verlag KGStuttgart · New York 0ce198Bilateral Saphenous Vein Thrombosis du male patient with multivessel coronary arterydisease who suffered from a nonST elevation a to admission a nonsmoker with type2 DM on metformin peptic ulcer diseaseand a history of CABG Just after the NSTEACS a left ventricle LV thrombus was seen on one echocar followup There was no deep venousthrombosis or bleeding diathesis history On admissi enoxaparin mg once daily Routine laboratory tests were withinnormal ranges ºTable There apart from obesity body mass index kgm2 when the patient was admitted The lower extren veins nosigns or symptoms of venous insufi¬ˆciency and the pastmedical history was negativ venous insufi¬ˆciency or varicose veins Thepatient was operated on following the star surgery resident harvested theright SV using the technique The wall of the SV lookedgrc were tiedoff and clipped and a needle was placed at the distal end whilethe proximal end was Results of initial and followup laboratory testingVariableCoagulation testsRed blood 103µLPlatelet count 103µLAPTT sPT sPT INRPlatelet aggregation mmolL arachidonic acid µ gLAntithrombin III Ddimer µgLantiXa IUmLHomocysteine µmolLProtein C Protein S Fac cï¬ˆbrinogen 455G AFactor XIII 100G TLupus anticoagulant ratioLupus anticoagul IgMAnti2glycoprotein I IgG antibodyAnti2glycoprotein I IgM antibodyNormal rangesPreope

0

no mutationGG no mutation GPL MPL SGU SMUGG no mutationGG no mutationGA

SMUAbbreviations APTT activated partial thromboplastin time GPL IgG phospholipid unit I
ratioMPL IgM phospholipid unit PT prothrombin time SGU standard IgG 2 glycoprotein unit S

0cBilateral Saphenous Vein Thrombosis during CABG Mazur et ale199to ï¬(ush the vein wit
normal saline mL while the distalend was closed with an atraumatic vascular clamp and vei
the distal end aluminal thrombus was visible The left SV was then taken downu
cardiacsurgeon with the same result Presence of a luminal thrombuswas conï¬^rmed upon s

was administered and normal LITA outï¬(ow wasconï¬^rmed Concerns regarding safety of

suspected thrombotic issueand the approach was modiï¬^ed The LITALAD anastomosis

postoperative course was uneventful On postoperative day the patient received dual antiplate
discharged on day with nosigns of thrombosis or myocardial ischemia Elective angiopla

completethrombophilia screening was done ºTable On the and12month followup the

thrombophilia was suspected screening wasinitiated showing no abnormalities ºTable

Positiveantibodies against neutrophil cytoplasm antigens pANCAand cANCA were excluded
analyze ï¬^brin phenotype using the previouslydescribed methodology34 Brieï¬(y plasm
hydrostatic pressure systemTubes containing ï¬^brin clots formed from adding mmolLcal
tocitrated plasma were connected through plastic tubing to abuffer reservoir M TrisHCl M Na
measured within minutes A permeation coefï¬^cient Ks reï¬(ecting poresize was calculated

Q is the ï¬(ow rate in time t L is the length of a ï¬^brin gelÎ· is the viscosity of liquid A is the cr

in dynecm2 Lower Ks values indicated reduced permeability Fibrinogen was determined usi
ï¬^brinogenlevel was normal we identiï¬^ed strongly decreased ï¬^brin clotpermeability K

from our previous report n ¼ Ks ¼ 9cm23 samples collected during late follow[] up appointme

controls n ¼ we observed prolonged clot lysistime CLT 06 vs 06 minutes and increased
thrombinpotential ETP in the studied subject ETP ¼ 06 vs 06 nM 02 min respectively mea
with calibrated automated thrombography thrombinoscope BV Maastricht theNetherlands a

96well plate ï¬(uorometer Ascent Reader Thermolabsystems OY Helsinki Finland equ

Brieï¬(y microliters of plateletpoor plasma were diluted with µL of the reagent contain
phosphatidylserinephosphatidylcholinephosphatidylethanolamine vesicle and µL
bovinealbumin and mmolL ZGlyGlyArg7amino4methylcoumarin Each plasma sample was
concentration of thrombin generated was used3Cryosectioned tissue sections were ï¬^xed

activity was quenchedwith H2O2 and unspeciï¬^c background was blocked with b
UnitedStates Primary adequate antibodies against ï¬^brin or tissuefactor TF were applied b

antibodies were followed by thecorresponding secondary antibodies conjugated with ï·
Images were analyzed using Olympus BX microscope SVs immunostainingreve

endotheliumºFig 1A and high TF ºFig 1B activity Within the thrombuswe found abunda

suggesting the presence of proinï¬(ammatory monocyteswhich are known source of TF
CD68 due to high unspeciï¬^c backgroundresulting from large amounts of ï¬^brin The mi

vessels ºFig 1C D Withinalmost every single vessel we found thrombi rich in bothprothromb

clot phenotype reï¬(ected by reducedKs and prolonged CLT along with enhanced thrombi
the immunostaining of the SVs prompted us to perform analysis ofwhole blood clot morpholo
previously described6 After washing thethrombus was ï¬^xed with glutaraldehyde phos
dehydrated goldcoated and photographed digitally with a JEOL JSM JEOL Tokyo Japan
layer on the clot surface with a solid massof unusually compressed platelets and erythr
veryhigh contractileforces during clot formation in a plateletdriven ï¬^brinmediated mechani

common ï¬^brinogen and factor F XIII polymorphisms The patient was heterozygous for ï¬^br
TDiscussionA dramatic intraoperative SV thrombosis provoked by graftharvesting for CABG
its cause remained unknown following thestandard thrombophilia screening The cases
period are very rareand as few as of grafts occlude within ï¬^rst to weeks17TH Vol No 0ce2

CABG Mazur et alFig Representative images of SV graft immunostaining after massive thro

green nucleistained blue using DAPI and scanning electron microscopic images E F of the

0

citrated bloodobtained from the patient undergoing CABG Box and arrow represent magniï¬
pertinent stainedfragments see text CABG Coronary artery bypass grafting SV saphenous v
of the intima the media and theadventitia8 The intima is built of the layer of endothelial cellsor
musclecells and the adventitia forms the outer part8 In a normalsetting the endothel
thrombosis9 and its focal disruption maypredispose to vessel thrombosis2 SV manipulat
integrity and elicitsan inï¬ammatory response with platelet adhesion and leukocyte re
isextremely rare in the operating room SV dissection results inblood ï¬ow disruption ir
hypoxia and vessel wall hyponutrition10 Acuteperioperative saphenous vein graft failure is al
very uncommonly occursprior to graft placement Surgical factorslike technicalanastomotic fa
vessel and the graft may lead to thrombosis butvessel injury and hypercoagulability are
evident inï¬ammatory process in microscopy inour patient but even if an inï¬amma
Saphenous Vein Thrombosis during CABG Mazur et ale201preoperatively in our patients S
not explain the dramatic intraoperativethrombosis We hypothesized that increased thro
phenotype were responsiblefor the clinical presentation Conversion of ï¬^brinogen to ï¬^br
coagulation Ithas been shown that ï¬^brin clots with small pores betweentight
resistant12Such clot phenotype has been evidenced in multiple thromboticpathologies such a
venous thromboembolism4 The prothrombotic clotphenotype reï¬ected by a tendency to for
previously reported in patients withinstent thrombosis14 While routine thrombophilia scr
commonhypercoagulable states15 there are prothrombotic conditionsthat escape routine di
appearance and prothrombotic ï¬^brin properties lead tothe discovery of two mutations i
thrombophilia screening namely ï¬^brinogen 455G A and FXIII100G TElevated ï¬^brinogen w
graft failure after CABG1116 Epidemiologicalstudies have established that elevated ï¬^brir
diseases17 A metaanalysis of individual records of participants from prospective studies rev
gL increase in usual ï¬^brinogen level forcoronary heart disease was conï¬^dence interval [(
as 95CI Risk of coronary disease progression wasalso linked to genetic polymorphisms of
allele of ï¬^brinogen 455G Awas associated with more severe progression of coronarydise
colleagues in a metaanalysis of studies with patientsfound that A allele of the ï¬^brinc
coronary disease and also withischemic stroke odds ratio for stroke ¼ [CI]for AA Ã¾
atrialï¬^brillation Hu and colleagues found that the A allele of ï¬^brinogen 455G A was a
elevating the level of plasma ï¬^brinogen20 On the other hand in a metaanalysis of stuc
polymorphism was shown to be associated withrisk myocardial infarction21 FXIII is cru
plasma concentration reï¬ectnonspeciï¬)cally the extent of thrombosis as shown by Li
thrombosis22Interesting associations of FXIII Val34Leu polymorphism andthrombotic disorc
metaanalysis of studies that FXIII Val34Leu polymorphism is associated with recurrent pre
incidence of ischemicstroke was found for this polymorphism24 apparentlywhen the stroke
the severity of its outcome25 Furthermore Kreutzand colleagues suggested in that FXI
recurrent MI and death inpatients with angiographically established coronary arterydisease2
that in patients undergoing CABG FXIII Leu34 alleleis associated with decre
ï¬^brinolysis27ConclusionOur extensive workup showed that ï¬^brinogen HaeIII andFXIII Val3
clotpermeability and susceptibility to lysis These mutationslikely contributed to intraoperative s
needed to elucidate the role ofthese polymorphisms in early graft failure after bypassgra
seemsevidentFundingThis study was funded by a grant from the JagiellonianUniversity Med
InterestNone declaredReferences Bourassa MG Fate of venous grafts the past the pres
Rosenfeldt FL Richards SM Conyers RA Davis BBImproved preservation of saphenous vei
solution during harvesting Circulation 19959209II31II36 Mazur P SokoÅowski G Hubal
Undas A Prothrombotic alterations in plasma ï¬^brinclot properties in thyroid disorders and
Undas A Zawilska K CieslaDul M et al Altered ï¬^brin clotstructurefunction in patients v
theirrelatives Blood Natorska J Marek G Hlawaty M Sadowski J Tracz W Undas AFibr
aorticstenosis association with in vivo thrombin generation and ï¬^brinclot properties Thromb
et al Reduced clot permeability and susceptibility to lysis in patients with acute coronarys
stress Atherosclerosis Fitzgibbon GM Kafka HP Leach AJ Keon WJ Hooper GD BurtonJR C
angiographicfollowup of grafts related to survival and reoperation in patients during years J

		0
		NJ Adams S Whellan DJ Saphenous vein graft disease review of pathophysiology prevent
		AW Endothelial cell injury in cardiovascularsurgery the intimal hyperplastic respon
		Prendergast FJ Veintoartery graftsthe longterm development of neointimal hyperplasia anc
		Thrombosis during CABG Mazur et alrelationship to vasa vasorum and sympathetic innerv
		Baisden CE de Winter RJ Alexander JHSaphenous vein graft failure after coronary artery byp
		future directions Ann Surg Undas A Fibrin clot properties and their modulation in thrombotic d
		Zawilska K et al Altered ï¬ˆbrin clotstructurefunction in patients with cryptogenic ischemic str
		Altered plasma ï¬ˆbrin clotproperties are associated with instent thrombosis ArterioscIerT
		KarpiÅ(ski M Stanisz A Undas AHigh detection rates of antithrombin deï¬ˆciency and antip
		years using thestandardized protocol for thrombophilia screening Thromb Res Moor E
		RydÃ©n LHaemostatic factors and inhibitors and coronary artery bypassgraftin
		occlusionThromb Haemost Danesh J Collins R Appleby P Peto R Association of ï¬ˆbrinoge
		with coronary heartdisease metaanalyses of prospective studies JAMA de Maat MP Kastele
		the betaï¬ˆbrinogen gene is associated with the progression ofcoronary atheroscl
		acutephase reaction pattern of ï¬ˆbrinogen REGRESS groupArterioscler Thromb Vas
		ï¬ˆbrinogen455GApolymorphism on development of ischemic stroke and coronaryheart dis
		ï¬ˆbrinogen gene 455GA polymorphism associated with cardioembolic stroke in atrial ï¬ˆ
		Jung JH Song GG Kim JH Seo YH Choi SJ Association of factor XIIIVal34Leu polymorphis
		Cardiol J Li B Heldner MR Arnold M et al Coagulation Factor XIIIin Cerebral Venous Throm
		Choi SJ Association of the F13A1Val34Leu polymorphism and recurrent pregnancy loss a m
		Shemirani AH PongrÃ¡cz E Antalï¬ˆ B AdÃ¡ny R Muszbek L FactorXI
		sufferingatherothrombotic ischemic stroke Thromb Res Shemirani AH Antalï¬ˆ B PongrÃ
		subunit Val34Leu polymorphism in fatal atherothrombotic ischemic stroke Blood Coagul Fibri
		XIII Val34Leu polymorphism and recurrent myocardialinfarction in patients withcoronary arte
		Plicner D Kapelak B Wypasek E Sadowski J UndasA Factor XIII Val34Leu polymorphisn
		resistance to lysis in patients withsevere coronary arte
3	Thyroid_Cancer	Solitary plasmacytoma SP of the skull is an uncommon clinical entity that is charac monoclonal plasma cells This case report describes a50yearold male that presented with occiputThe diagnosis of SP was based on the pathological results and imaging examinat skull reconstruction and lower trapezius myocutaneous flapLTMF transplantation under gener extended tothe subcutaneous and the subdural space through the dura mater with skull defec large areas of scalp and subcutaneous tissue which resultedin a large postopera transplantation All of thetumour was removed and the transplanted flap grew well Followup on the right frontallobe The patient received six cycles of the PADchemotherapy regimen bc the lesion was significantly reduced This case demonstrates that LTMF is an alternative ap soft tissue defects caused by the excision of a large malignant tumourof the occipital regi neoplastic recurrenceKeywordsSolitary plasmacytoma lower trapezius myocutaneous flap July accepted March 1Department of Neurosurgery Hunan Cancer Hospitaland the Affiliated Central South University Changsha HunanProvince China2Department of Head and Neck S Cancer Hospital of XiangyaSchool of Medicine Central South University ChangshaHunan P and ZhengWen He Department ofNeurosurgery Hunan Cancer Hospital and the Affiliated CentralSouth University Tongzipo Road Yuelu DistrictChangsha Hu hezhw2001163comCreative Commons Non Commercial CC BYNC This is distrib AttributionNonCommercial License creativecommonslicensesbync40 which permitsnoncom work without further permission provided the original work is attributedas specified on the SA(accessatsage 0cIntroductionSolitary plasmacytoma SP is the pathological manifestation of an SPthat originates in bone tissue is called a solitary plasmacytoma of bone SPB1 Bonedes the most common sites are thepelvis spine femur humerus and ribs2An SPB of the skull is

0
rarely mentioned in the literature34 Complete tumourremoval is the ï¬ˆrst and best apprc
thebody3 This current case report describes arare case of SPB of the occipital bone withsc radical resection and reconstruction ofJournal of International Medical
trapeziusmyocutaneous ï¬(ap LTMF transplantationCase reportA 50yearold male ofNeurosurgery Hunan Cancer Hospitaland the Afï¬ˆliated Cancer Hospital ofXiang
Changsha HunanProvince China with a headache and an exophytic mass on the occipu
Computed tomographyCT showed a large mass with homogeneousenhancement on the c and theneurologicalFigure Preoperative imaging examinations a preoperative appearance of
tomography scan c preoperative enhanced magnetic resonance imaging MRId preoperative
digital subtraction angiographyThe colour version of this figure is available at httpimrsagepu revealed a solitary osteolyticlesion involving the whole entire of the occipital b extracerebralexpansile osseous lesion 02 mm themass was mostly isointe T2weightedimagesenhancedFigures 1c and 1d Digital subtraction angiograp hypervascularity that was supplied from the occipital artery In order todecrease bleeding v blood vessel wasembolized during DSA ehomogeneouslyThe patient underwent occ transplantation under general anaesthesia Thetumour was capsulized and extended tothe sul dura mater with skull defectsGrossly the tumour had a ï¬ˆshmeat likeappearance mixed wi
had arich blood supply and despite embolizationof the main blood supply artery duringD during the operation Thetumour mass underwent extended resectionincluding the marginal I windowand a 02 cm scalp defect The skulldefect was reconstructed using titaniummesh an The trapezius and the skinisland 02 cm and the supplying vesselsof the transverse ce
marked on theskin b The island ï¬(ap was excisedand its muscle pedicle dissected up to th examinations a the tumour was fishmeat soft tan inappearance b the trapezius and transverse cervicalartery and the dorsal scapular artery marked out on the skin c the islan flap was set into the defect with a wellperfused distal end e the stiches were removedat
resonance imaging MRI scan sagittal viewg postoperative enhanced MRI scan axial view
httpimrsagepubcom 0crotation point at the medialsuperior edgeof the scapula c The LTMF wa through the neck posterior subcutaneous tunnel d Two weeks afterthe operation the trans undisturbedFigurethattheremovedtumour wasFigures 2f and 2g2e MRIcompletelyindicate
showed the presence of atypicalplasma cells with typical eccentric roundnucleistainingsho
epithelial membrane antigen melanA CD38 ¾ CD138ImmunohistochemicalJournal of Int
Lambda ¾glial ï¬ˆbrillary acidic protein S100CD68 ¾ thyroid transcription factor1 Vim CI
radiotherapyfor ï¬ˆnancial reasons After a followupperiod of around months he was symptom
the 5month followup visitMRI revealed no ˆeld recurrence butan aggressive mass lesio
Chemotherapy PADregimen bortezomib pegylated liposomaldoxorubicindexamethas thefrontalrightandFigure Representative photomicrographs of the tumour a haematoxylin and plasma cells b immunohistochemical staining for CD138 showing strong positivity in the tu CD38 showing strong positivity in thetumour cells d the positive expression of Ki67 w athttpimrsagepubcom Scale bar mm 0cWang et alFigure Magnetic resonance imaging scans followup visit showing no recurrence in situ but an aggressive mass lesion with er consecutive cycles of chemotherapy showing no recurrence in situ and the
reducedDepartmentof Haematology HunanCancer Hospital and the Afï¬ˆliated CancerHosp
University ChangshaHunan Province China After six consecutive cycles of chem signiï¬ˆcantly reducedFigures 4c and 4d Postoperative reviewafter months showed no tum
was a case report the InstitutionalReview Board of Hunan Cancer Hospitalwaived the need informed consentfor publication that was approved by theInstitutional Reviev anonymizedDiscussionHuge intra and extracranial SPs of theoccipital bone are very ra characterizedby the presence of a solitary lytic lesion dueto monoclonal plasma
withoutaccount for of all SP cases and theyoccur primarily in red marrowcontainingbones6 Radiation therapy remains the 0cJournal of International Medical Researchtreatment
recommendations froma European expert panel a total fractionated dose of Gy sh
beemployed6 In this current caseit wasunfortunate that the patient refused radiationre operation and a new mass was foundon the right frontal lobe After six cyclesof chemothera
which suggests that chemotherapy has a positive impact on thegrowth of recurrent tumourse of SPB in theskull are complex and can easily lead tomisdiagnosis Enhanced CT s windoware credible means by which to diagnoseSPB and they could provide more informa the MRI examination allowed for theidentiï¬ˆcation of the location size and shapeof the tum
structures In our opinionpreoperative DSA is necessary for the identiï¬ˆcation of

		0
		embolizationDuringthetumourrecurrencecurrent operationthetumour was found to involv theprobability oftheinvolved scalp underwent an extendedresection LTMF was used to facilita available muscle compartments transferred on a reliable vascular pedicle to thedorsal sup LTMF include well vascularized tissue ease of harvest and the provision of a large ï¬(ap loca The mainblood supply to the LTMF originates fromthe transverse cervical artery and the dor the problem ofinsufï¬^cient bloodsupply caused by titanium plate implantation In addit subcutaneous cavitycreated by the huge tumourresectionpreventing occipitalia scalp hydrops approach for the repair of scalp andsubcutaneous soft tissue defects caused byexcision of contributionsLW studied the case collected the referencesand wrote the paper ZH designed t thepaper XP analysed the data NR served as theï¬^rst chief during surgery and wr ï¬^nalmanuscriptDeclaration of conflicting interestThe authors declare that there are supported by grants from theScientiï¬^c Research Project of Hunan ProvincialHealth Commis Science Foundation of ChinaNo2019JJ40182 and the Sailing Programme ofHunan Provinc iDLei WangReferences Sabattini E Bacci F Sagramoso C et alWHO classiï¬^cation of tum anoverview Pathologica Gee ED and Sadovsky R Multiple myeloma recognition and manage Ghehit¸ 15a KLet al Neurosurgical rare disease solitary 0cWang et alplasmacytoma of th Morphol Embryol Chang MY Shih LY Dunn P et al Solitaryplasmacytoma of bone J Fo extramedullary plasmacytoma Hematology Am SocHematol Educ Program Caers J Pa response assessment in solitary plasmacytoma updated recommendations from a Eu Zouhair A Tsang RW et alPrognostic factors in solitary plasmacytomaof the bone a multic Liebross RH Ha CS Cox JD et alSolitary bone plasmacytoma outcome andprognostic facto Phys Mohos G Vass G Kemeny L et alExtended lowertrapezius myocutaneousï¬(ap to cove new application J PlastSurg Hand Surg U 15gurlu K Ozc¸ elik D Hu¨ thu¨ t I et alExtended and neck reconstruction as asalvage procedure Plast Reconstr Surg Baek SM Bille myocutaneous ï¬(apAnn Plast Surg Netterville JL and Wood DE The lower trapezius ï¬(ap
4	Thyroid_Cancer	This study aimed to investigate serum matrix metalloproteinase MMP2 and MMP9le PTCMethods Fortyone patients with PTC undergoing ultrasoundguided radiofrequency MMP2 and MMP9 levels were determined byenzymelinked immunosorbent assay bef wereevaluated by logistic regression analysisResults Serum MMP2 and MMP9 comparedwith controls and decreased significantly after surgery According to receiver op values for preoperative serum MMP2 and MMP9 levels were and There was no contrastagen enhancement within or at the lesion edge in The volume reduction at monthsfollowup was A diameter and numberwere influencing factors for PTC Age and lesion diameter and number and morphology were protective factorsConclusion Serum MMP2 and MMP9 levels andtreatment of PTC by RFA Preoperative serum MMP2 and MMP9 levels combined prognosisDepartment of Ultrasound Beilun Peoples Hospital ofNingbo Beilun Branch of th Ningbo Zhejiang ChinaThese authors contributed equally to this studyCorresponding aut PeoplesHospital of Ningbo Beilun Branch of the First AffiliatedHospital of Zhejiang Un ChinaEmail qianding02sohucomCreative Commons Non Commercial CC BYNC This is distrib AttributionNonCommercial License creativecommonslicensesbync40 which permitsnoncom work without further permission provided the original work is attributedas specified on the SA(accessatsage 0cJournal of International Medical ResearchKeywordsPapillary th metalloproteinase2 matrix metalloproteinase9 disease prognosis serum levelDate recei thyroid carcinoma PTCis acommon thyroid malignancy accountingfor about of systemic ma highincidence and low mortality2clinicalRecentimprovementsin highfrequencyultrasoni ultrasoundguided puncturetechniques have led to an apparent increasein the incidence of palpation detection rate forthyroid microtumors in the general population is about compared which has thus greatly improveddisease diagnosisdiagnosistheofSurgical resection recurrence rateis usually high and the consequent reduction in thyroid function can seriously a the increasing detection rate of thyroidtumors and the pursuit of minimally invasive tre

0

gradually applied in theclinic RFA uses local hyperthermia tocause tissue necrosis The t
andmost lesions can be completely eliminatedby RFA89 RFA has thus become a novelloc
benign andmalignant thyroid tumors currently dependson the clinical manifestati
clinicalmanifestations are mostly derived frominvolvessubjective empirical analysis while a ï
with lesssatisfactory speciï¬ˆcity It is therefore necessary to identifyappropriatepredictive
collagenases matrix metalloproteinase MMP2 and MMP9 can degradety
importanttumorangiogenesis and tumor cell invasion andmetastasis10 MMP2 and MMP9 ex
thyroid cancer tissue11 however thesestudies mostly examined pathological tissuesafterinvas
serum levels of MMP2and MMP9 have been lesswell consideredIn this study we detected se
PTCbefore and after ultrasoundguided RFAWe also determined the therapeutic effectsof RFA
relevant prognostic factorsMaterials and methodsStudy subjectsPatients who underwent
toOctober were included in this studyThe inclusion criteria were as follows patients dia
cytology no historyof neck surgery and patients requiringminimally invasive treatment for aes
0cPan et alwith anxiety The exclusion criteria were asfollows benign lesions conï¬ˆrmed by
surgery and severe coagulopathyPeripheral venous blood samples wereobtained from the i
the operation and serum levels of MMP2 andMMP9 were determined Additional subjects
RFA were included as a controlgroup Prior written informed consent wasobtained from all pa
review boardof our hospitalPreoperative preparationcalciï¬ˆcationThe number size nature
nodular blood ï¬ow distributionof the tumors were assessed before the operation After skin d
lidocainesolution A total of mL Sonovue Bracco Milan Italy was injected via the elbowvei
evaluated by contrastenhancedultrasound CEUS of the ablationtargetedlesions using a Myla
ShenzhenGuangdong ChinaAccording to the location of the thyroidnodules the thyroid and c
and esophageal space and posterior thyroid spacerecurrent laryngeal nerve were separate
the intraoperative conditions to form aliquid separation zone to protectthesestructures
ultrasound guidancethe tip ofthe RFA needle rated power Woutput frequency kHz was
wasperformed using an OlympusCelon PowerRFA System Germany in mobile mode12followi
were subjected to multipointed and multifaceted ablation untilthe thyroid tissue layer with the
echogenerated by heat accumulation The wholeprocess was carried out under continu
produced in the ablation zone duringthe ablation treatment The position of theelectrode
lesion size After ensuringthat there was no residual enhancement inthe ablation zone the abl
completedAfter the operation an ice compress wasapplied for h to avoid skin burnsSerum
blood wascollected from the elbow vein under fastingconditions before and after
controlgroup were collected after ultrasound contrast examination The blood samples werep
subjected to centrifugation at 02 g for minutes The serum washarvested and serum lev
enzymelinked immunosorbentELISAkitsBoster Bioengineering WuhanHubei ChinaassayFol
the operation the ablation range was evaluated by CEUS If residualtissues were detected
detectionwas performed at and monthsafter surgery to determine the nodule sizesand volum
according to the followingformula volume reduction rate¼ preoperative volume foll
bloodï¬ owchanges in the ablation zone were alsoobserved and analyzed The efï¬ˆcacy
treating tumors13 disappearance ofnodules indicated by complete disappearance of blood
complete cure noduleby 15 indicatedvolumemarked ef
improvementreducedClinicopathological featuresInformation on ultrasoundbased clinicopath
of the lesion wereobtainedfeaturesversion Statistical analysisData were expressed as mea
performed using IBM SPSS StatisticsforWindowsIBM CorpArmonk NY USA Com
v2testsPotentially related factors were analyzedby univariate or multivariate logistic regres
MMP2 and MMP9 levels wereevaluated by receiver operating characteristic RC
signiï¬ˆcantResultsPatientsJournal of International Medical Research men mean age 06 yea
beforeand after treatmentThe characteristics of the ultrasound images inthe included subj
and MMP9 weremeasured before and after treatment Serumlevels of MMP2 and MMP9
comparedwith the control subjects P Serumlevels of MMP2 and MMP9 had declinedat
surgery but the difference wasnot signiï¬ˆcant However serum levels ofMMP2 and MMP9 ha
and monthsall P Table These results suggestthat changes in serum MMP2 and MMP9lev
signiï¬ˆcant implications for the therapyof PTCROC curve analysis of preoperative serumMM
and MMP9levels were used as potential diagnostic indicators In the patients with PTC the
was used as the diagnostic resultsand the gold standard classiï¬ˆcation criteriawere us
obtained accordingly The areaundercurve AUC values for serum MMP and MMP9 lev
suggestthat serum levels of MMP2 and MMP9could contribute to the disease dia
enrolledincluding women and men mean age 06 years range to Ã¾65 years The
benignthyroid nodules including women andEvaluation of RFA efficacyWe also evaluated t
ablation showed hypoenhancement in nodules isoenhancementin nodules and slight hyperer
thyroid ultrasound imagesPTC patientsNormal controlLe
cmCalcificationMicrocalcificationCoarse calcificationMorphologyRegularIrregularA

0

groupnTable Serum matrix metalloproteinase2 and levels in controls and inpatients w
treatmentControlsPTC patientsBefore surgery month after surgery months after surgery mont

06 06 06 06 06 PMMP9P 06 06 06 06 06 06 MMP matrix metalloproteinase PTC pa

examination afterablation showed no contrastagent perfusionin the ablation zone in le
theedge or inside the lesion in the other lesions and the ablation area wasgradually reduc
signiï¬^cant changein ablation volume in any patients at month after surgery compare

reducedby at months offollowupcompared with before surgery P Table These results showed
with PTCcould effectively reduceInfluence of relevant factors on diseaseprognc
retrospectively analyzed by 0cJournal of International Medical Researchirregularlogistic

disease prognosis Agemicrocalciï¬^cationshape anddiameter and number of lesions were

hazard ratios HRsfor age andlesion diameter and number were indicating that these repr
HRs for microcalciï¬^cation and irregular shape were negativeindicating that a greater

associated with lower risks of developing the disease and were thus protective facto
andMMP9 levels MMP matrix metalloproteinaseFigure Efficacy evaluation of radiofrequ
showingobvious bloodflow signals around the tumor and fewer signals within the tum
showing no obvious enhancement in the lesion with lowperfusion performance c Inthe 2
inserted into the tumor for ablation d The tumor wascompletely ablated with no

alDiscussionitis difi¬^cultPTC is a type of thyroid tumor with a highincidence14 wl

Mostthyroidtumor cases are currently diagnosed by hiscytological detectiontopathologic
malignant papillaryhyperplastic nodules and it is therefore difi¬^cult to diagnose PT

prognostic molecular markers for PTC17 The relationshipbetween MMPs and tumors
MMPsplay important roles in pathophysiologicalprocesses such as the dynamic extracellula
and repair10 Tumor cells may inducethe matrix to secrete MMPs via a series ofsignalingp
after radiofrequencyablation of papillary thyroid carcinomasReductionrate Lesionvolum

compared with before surgeryBefore surgery month after surgery months after surgery mor

cell invasionand metastasisalternatingthe extensive surgicalSurgical resection is a traditi
ofHowevertraumaunsightly neck scars and risks of laryngealnerve injuries postoperativ
increasing numbers of patients are opting for minimally invasive ablation meth
highfrequencyelectromagneticwaves generated by the radiofrequencyinstrument inserted ir
friction of positive and negative ions within the cells causing local coagulation

bodyssystem19 Reduction ratesimmuneforbenign thyroid nodules of to after month

beenreported20 In this study ultrasound performed immediately after ablation of lesions sh
and enhancementto varying degrees at the edge or inside thelesion in of lesions Althou

volumeat month after the operation the lesionvolumes were reduced by to at months after sur
be associatedwith the heat and the ablation needle aTable Logistic regress
prognosisAgeMicrocalcificationIrregular morphologyLesion diameterLesion numberB HRHI
limitUpper limitP 0cJournal of International Medical Researchï¬^ne needle is good

alternatingelectromagnetic wave only circulates in theeffective region between the two
ablation zone The ablation safety zone aroundthe PTC was relatively small in the currents
was thus relatively greaterWe analyzed the serum levels of MMP2and MMP9 in PTC patie
enzymeswere signiï¬^cantly higher in patients withPTC compared with patients with ben

MMP2 and MMP9 levelstheAUC values based on the ROC curveswere and for MMP2
clinical diagnostic and prognostic valuesIn the present study serum levels MMP2and MM

compared with before surgerybut the difference was not signiï¬^cant Thismight be becaus

the tumorstroma to produce larger amounts ofMMP2 and MMP9 which were releasedint
completely ablate the tumor ortumor recurrence may result in the secretionof high levels of M
levels of MMP and MMP9 were signiï¬^cantly decreasedat and months after surgery comp

MMP2 and MMP9 were secretedby the tumor The lesions disappearedafter PTC ablation th
and therebyreducing the degradation and destructionof type IV collagen protecting the base
the growth and metastasis oftumorresults alsoshowed that age microcalciï¬^cation irregu

Thecurrentwere risk factors for PTC Logistic regression analysis showed that age 14 years
the ï¬^ndings of Yu et al21Microcalciï¬^cation is caused by the deposition of calcium salts a

salts by thetumor itself and has been considered tobe the most speciï¬^c sign of PTC In th

higher in PTC patients compared withthe control group and logistic regressionidentiï¬^ed it

results also identiï¬^edirregular morphology two nodules and anodule diameter 14 cm

previousï¬^ndings22This study had some limitations It was asinglecenter study with a

relevantthyroid hormone analysis and other serumindicators could not be followed up

summary the results of this studyshowed that RFA could shrink or eliminatethyroid lesions t
effective methodforserumlevels of MMP2 and MMP9 before RFAcould provide a val
serological indexes combined with relevant risk factors mayalso help to predict the prognos

	0
	work wasProvincial HealthCommission Project WJ2017F102supported by the HubeiPlanninga
	authors declare that there is no conï¬(ict ofinterest 0cPan et alFundingThis research receive
	the public commercial ornotforproï¬^t sectorsorcid000000026050ORCID iDQian DingRefer
	compartment reoperation for recurrentpersistent differentiated thyroid cancer]Zhonghua Er E
	Pellegriti G Frasca F Regalbuto C et alWorldwide increasing incidence of thyroidcancer up
	Epidemiol DOI Brito JP Gionfriddo MR Al Nofal A et alThe accuracy of thyroid nodule
	reviewand metaanalysis J Clin Endocrinol Metab DOI 101210jc2013 Zhao P Zheng D
	thyroid microcarcinoma a report of cases Chinese JGene Surg Anil G Hegde A and C
	malignancywith ultrasound and guided biopsy CancerImaging DOI Levine RA Current gu
	Endocr Pract DOI 104158ep12071co Zhang Y Zhang MB Luo YK et al Effectof chronic lymp
	ultrasoundguidedradiofrequency ablation for papillary thyroid microcarcinoma Cancer Med
	percutaneous microwave ablation inthe treatment of recurrent thyroid nodulesJ Clin Otolaryng
	Mourao GRegina Calsolari M et al Role of adjuvanttherapy with radioactive iodine in patie
	reoperation due to recurrent papillarythyroid cancer a monoinstitutional comparative s
	MMP2MMP9 TIMP1 and TIMP2 in theperipheral blood of patients with differentiated thyr
	Tang J et al Ultrasoundguidedradiofrequency ablation for papillary thyroidmicrocar
	Hyperthermia Zhao CK Hu HX Lu F et al Factors associated with initial incc
	radiofrequencyablation First results of CEUS evaluationClin Hemorheol Microcirc Faggiano
	treated with percutaneousradiofrequency thermal ablation a comparative study J Clin Endo
	EJ Cachia AJ Chung HR et alResveratrol supplementation reduces aorticatherosclerosis a
	capacity in a mousemodel of uremia J Med Food DOI 101089jmf20120219 Xhaard C Ru
	factors in therisk of differentiated thyroid carcinoma inyoung women in France a populatio
	101093ajekwu220 Tafani M De Santis E Coppola L et alBridging hypoxia inï¬(amr
	progressionBiomed Pharmacother DOI101016jbiopha201310013 Bumber B Marjanovi
	metalloproteinases and their inhibitors in the development of cervical metastase
	Researchthyroid cancer Clin Otolaryngol ofAssociation Gharib H Papini E Paschke R et al
	MediciEndocrinologiand European ThyroidAssociation medical guidelines for clinicalpracti
	nodules executive summary ofrecommendations J Endocrinol Invest Baek JH LeeJH V
	nodules radiofrequency and laser Korean JRadiol DOI kjr2011125525 Vuong NL Dinh
	nodules 1yearfollowup in patients World J Surg Lang BHH Woo YC and Chiu KWIdentifying
	focused ultrasound HIFUablation of benign thyroid nodules a retrospective analysis Int J Hyp
	Canmalignant thyroid nodules be distinguishedfrom benign thyroid nodules in children andac
	pe

In [7]:
```
#Rename the column names
data.rename({'0':'Label','a':'Text'},axis=1,inplace=True)
```

EDA Process

In [8]:
```
#Check the data columns
data.columns
```

Out[8]:
```
Index(['Label', 'Text'], dtype='object')
```

In [9]:
```
data[data.Label=='Lung_Cancer'].Text[539]
```

Out[9]: '"We lack biomarkers for identifying aggressive primary tumor subsets that giv e rise to metastases and impact early cancer detection and treatment. Many sol id tumors are known to accumulate hyaluronan (HA) a glycosaminoglycan which is also produced by the tumor cells themselves. We report a quantitative approach for uncovering breast cancer heterogeneity using fluorescent HA to detect diff erential binding patterns to CD44 and RHAMM/HMMR receptors. This approach perm its identification of tumor-cell subsets that bind high levels of HA and may b e applicable to other ligands/receptors and disease models. Despite representi ng the invasive/metastatic subset of parental tumors unexpectedly the high HA- binding subset was slow-growing and is thus likely to be a source of dormancy and relapse. Tumor heterogeneity confounds cancer diagnosis and the outcome of therapy necessitating analysis of tumor cell subsets within the tumor mass. El evated expression of hyaluronan (HA) and HA receptors receptor for HA-mediated motility (RHAMM)/HA-mediated motility receptor and cluster designation 44 (CD4 4) in breast tumors correlates with poor outcome. We hypothesized that a probe for detecting HA\x93HA receptor interactions may reveal breast cancer (BCa) ce ll heterogeneity relevant to tumor progression. A fluorescent HA (F-HA) probe containing a mixture of polymer sizes typical of tumor microenvironments (10\x 93480 kDa) multiplexed profiling and flow cytometry were used to monitor HA bi nding to BCa cell lines of different molecular subtypes. Formulae were develop ed to quantify binding heterogeneity and to measure invasion in vivo. Two subs ets exhibiting differential binding (HA?/low vs. HAhigh) were isolated and cha racterized for morphology growth and invasion in culture and as xenografts in vivo. F-HA\x93binding amounts and degree of heterogeneity varied with BCa subt ype were highest in the malignant basal-like cell lines and decreased upon rev ersion to a nonmalignant phenotype. Binding amounts correlated with CD44 and R HAMM displayed but binding heterogeneity appeared to arise from a differential ability of HA receptor-positive subpopulations to interact with F-HA. HAhigh s ubpopulations exhibited significantly higher local invasion and lung micrometa stases but unexpectedly lower proliferation than either unsorted parental cell s or the HA?/low subpopulation. Querying F-HA binding to aggressive tumor cell s reveals a previously undetected form of heterogeneity that predicts invasiv e/metastatic behavior and that may aid both early identification of cancer pat ients susceptible to metastasis and detection/therapy of invasive BCa subpopul ations. tumor cell heterogeneity hyaluronan binding heterogeneity index PLoS O ne one 1932-6203 Public Library of Science San Francisco USA 24454921 38932 58 PONE-D-13-41069 .0085702 Research Biology Biochemistry Bioenergetics Energ y-Producing Processes Metabolism Carbohydrate Metabolism Metabolic Pathways Ox ygen Metabolism Protein Metabolism Cofactors Drug Discovery Enzymes Genetics G ene Expression Medicine Drugs and Devices Drug Research and Development Drug D iscovery Hematology Hematologic Cancers and Related Disorders Leukemias Acute Lymphoblastic Leukemia Nutrition Obstetrics and Gynecology Breast Cancer Oncol ogy Cancers and Neoplasms Hematologic Cancers and Related Disorders Leukemias Breast Tumors Oncology Agents Metabolic Effects of Acute Thiamine Depletion Ar e Reversed by Rapamycin in Breast and Leukemia Cells Thiamine Depletion and Me tabolism in Cancer Cells Liu Shuqian 1 Miriyala Sumitra 2 Keaton Mignon A. 3 J ordan Craig T. 4 Wiedl Christina 5 Clair Daret K. St. 2 Moscow Jeffrey A. 1 * 1 Department of Pediatrics University of Kentucky College of Medicine Lexingto n Kentucky United States of America 2 Graduate Center for Toxicology Universit y of Kentucky College of Medicine Lexington Kentucky United States of America 3 Metabolon Inc Durham North Carolina United States of America 4 Division of H ematology Hematologic Malignancies and Stem Cell Transplantation University of Colorado Denver Colorado United States of America 5 Department of Pediatrics V irginia Commonwealth University Richmond Virginia United States of America Ahm ad Aamir Editor Wayne State University School of Medicine United States of Ame rica * E-mail: jmoscowuky.edu Competing Interests: One of the authors of this paper Mignon A. Keaton was employed by Metabolon Inc. during the data acquisit ion and analysis phases of the study. Dr. Keaton is no longer employed by Meta bolon. Her employment history does not alter the authors\' adherence to all th e PLOS ONE policies on sharing data and materials. Conceived and designed the

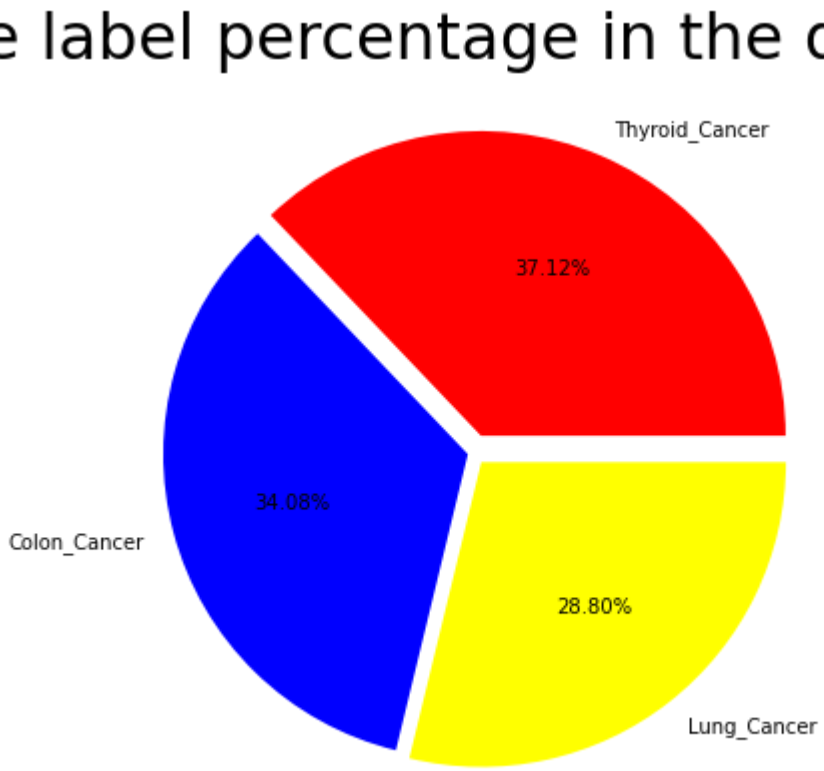
experiments: MAK CTJ DKS JAM. Performed the experiments: SL SM MAK CTJ CW. Ana lyzed the data: MAK DKS JAM. Contributed reagents/materials/analysis tools: SM MAK CTJ CW DKS JAM. Wrote the paper: MAK CTJ JAM. 2014 15 1 2014 9 1 e85702 8 10 2013 5 12 2013 2014 Liu et al This is an open-access distributed under the terms of the Creative Commons Attribution License which permits unrestricted u se distribution and reproduction in any medium provided the original author an d source are credited. Thiamine-dependent enzymes (TDEs) control metabolic pat hways that are frequently altered in cancer and therefore present cancer-relev ant targets. We have previously shown that the recombinant enzyme thiaminase c leaves and depletes intracellular thiamine has growth inhibitory activity agai nst leukemia and breast cancer cell lines and that its growth inhibitory effec ts were reversed in leukemia cell lines by rapamycin. Now we first show furthe r evidence of thiaminase therapeutic potential by demonstrating its activity a gainst breast and leukemia xenografts and against a primary leukemia xenograf t. We therefore further explored the metabolic effects of thiaminase in combin ation with rapamycin in leukemia and breast cell lines. Thiaminase decreased o xygen consumption rate and increased extracellular acidification rate consiste nt with the inhibitory effect of acute thiamine depletion on the activity of t he TDEs pyruvate dehydrogenase and 2-oxoglutarate dehydrogenase complexes; the se effects were reversed by rapamycin. Metabolomic studies demonstrated intrac ellular thiamine depletion and the presence of the thiazole cleavage product i n thiaminase-treated cells providing validation of the experimental procedure s. Accumulation of ribose and ribulose in both cell lines support the thiamina se-mediated suppression of the TDE transketolase. Interestingly thiaminase sup pression of another TDE branched chain amino ketoacid dehydrogenase (BCKDH) sh owed very different patterns in the two cell lines: in RS4 leukemia cells it l ed to an increase in BCKDH substrates and in MCF-7 breast cancer cells it led to a decrease in BCKDH products. Immunoblot analyses showed corresponding diff erences in expression of BCKDH pathway enzymes and partial protection of thiam inase growth inhibition by gabapentin indicated that BCKDH inhibition may be a mechanism of thiaminase-mediated toxicity. Surprisingly most of thiaminase-med iated metabolomic effects were also reversed by rapamycin. Thus these studies demonstrate that acute intracellular thiamine depletion by recombinant thiamin ase results in metabolic changes in thiamine-dependent metabolism and demonstr ate a previously unrecognized role of mTOR signaling in the regulation of thia mine-dependent metabolism. No current external funding sources for this study. Introduction Thiamine (vitamin B1) is a cofactor for enzymes involved in criti cal metabolic processes involving energy production biomass generation and ami no acid catabolism. Despite the requirement for this vitamin in these central processes the role of thiamine and thiamine-dependent enzymes (TDEs) in cancer development and treatment has received little attention although a recent revi ew has summarized the potential importance of TDE\x99s in cancer metabolism [1]. Unlike antifolates which have a well-established role in cancer therapy a nalogous small molecule thiamine antagonists are relatively inert leading to a that TDE pathways could not be important as an anticancer targets. However the limitations of small molecule TDE inhibitors should not be confused with the p otential role of TDEs as anticancer therapeutic targets. Antifolates can be ef fective because intracellular folates only transiently associate with enzymes during the catalytic process allowing for inhibition of enzyme activity by mol ecules designed to bind more tightly than the intracellular substrates. In con trast intracellular thiamine activated by phosphorylation remains tightly boun d to enzyme complexes during the catalytic cycle leaving little opportunity fo r inhibitors to displace it once the complex has assembled. This inherent phar macologic challenge could disguise the potential of targeting TDEs for cancer therapy. We have previously shown down-regulation of thiamine transporter gene expression in tumors compared to normal tissues [2] [3] and more recently have shown that a low thiamine diet delays onset of mammary tumors in MMTV(her2) mi ce [4] an effect that is abrogated by a high fat diet. These observations have led to our hypothesis that TDE pathways are altered as part of the overall cha nges in energy metabolism that occurs in cancer cells and that these changes c

```
ould produce metabolic vulnerabilities that could be exploited by therapies ai
med at TDE activities. To take a novel path in the exploration of TDEs in canc
er we have studied the cytotoxic activity of the bacterial enzyme thiaminase w
hich cleaves thiamine into its pyrimidine and thiazole moieties [5]. Thiaminas
e overcomes the limitations of small molecule TDE "'
```

In [10]:
```
#Let's count the label values in the dataset
label=data['Label'].value_counts()
label
```

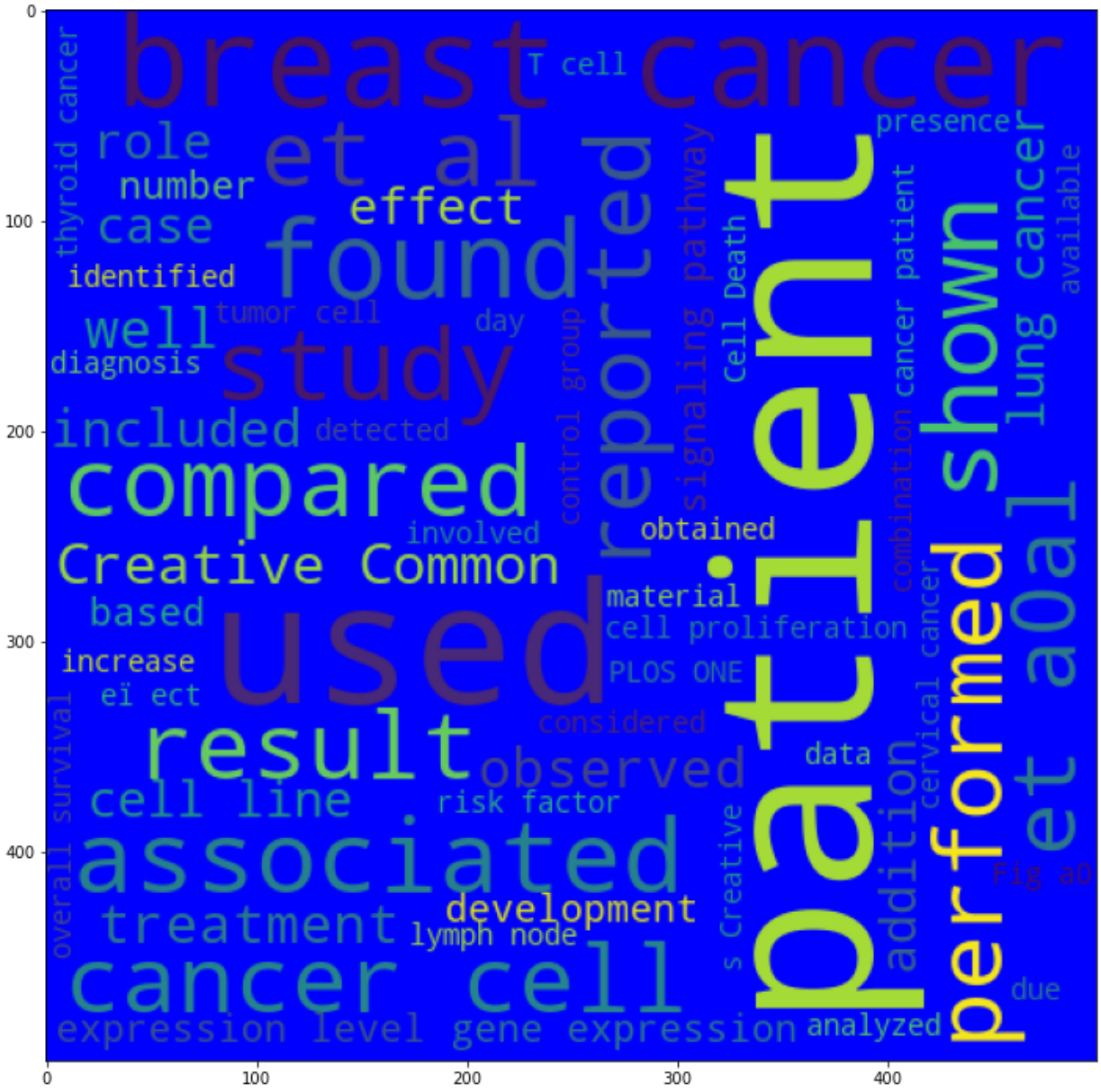
Out[10]:
```
Thyroid_Cancer    2810
Colon_Cancer      2580
Lung_Cancer       2180
Name: Label, dtype: int64
```

In [11]:
```
#Let's visualize the above information on the dataset
plt.figure(figsize=(19,7))
plt.pie(label,
    labels=['Thyroid_Cancer','Colon_Cancer','Lung_Cancer'],
    colors=['Red','blue','yellow'],
    autopct='%1.2f%%',explode=[0.06,0.02,0.04])
plt.title("The label percentage in the dataset",fontsize=32)
plt.show()
```



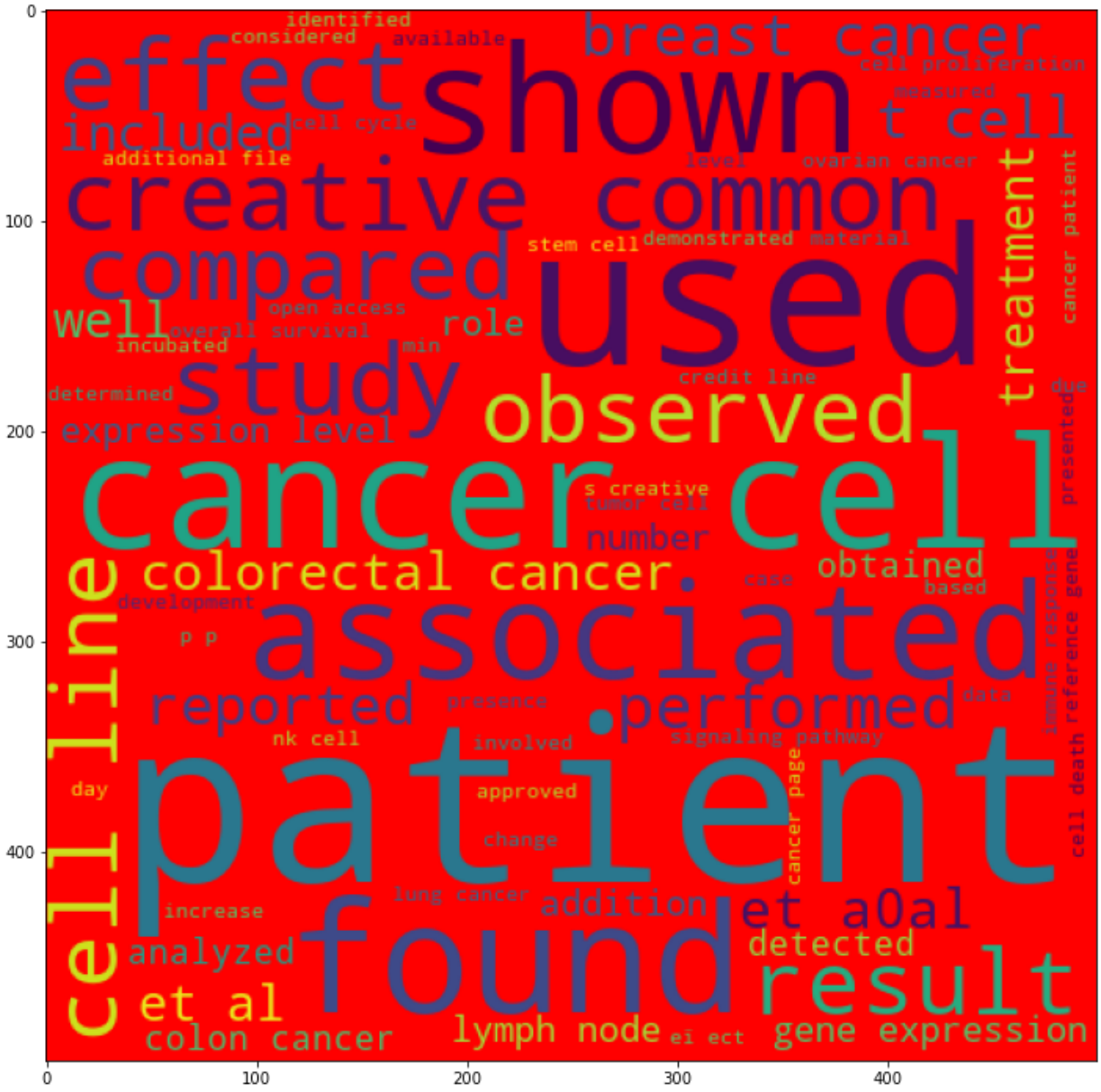
In [12]:
```
#Visualize the WordCloud in the Thyroid_Cancer from the dataset
plt.figure(figsize=(12,15))
wc=WordCloud(height=500,width=500,min_font_size=10,background_color='blue')
w_c=wc.generate(data[data['Label']=='Thyroid_Cancer']['Text'].str.cat(sep=" "))
plt.imshow(w_c)
```

Out[12]:
```
<matplotlib.image.AxesImage at 0x7fc925a8b5e0>
```



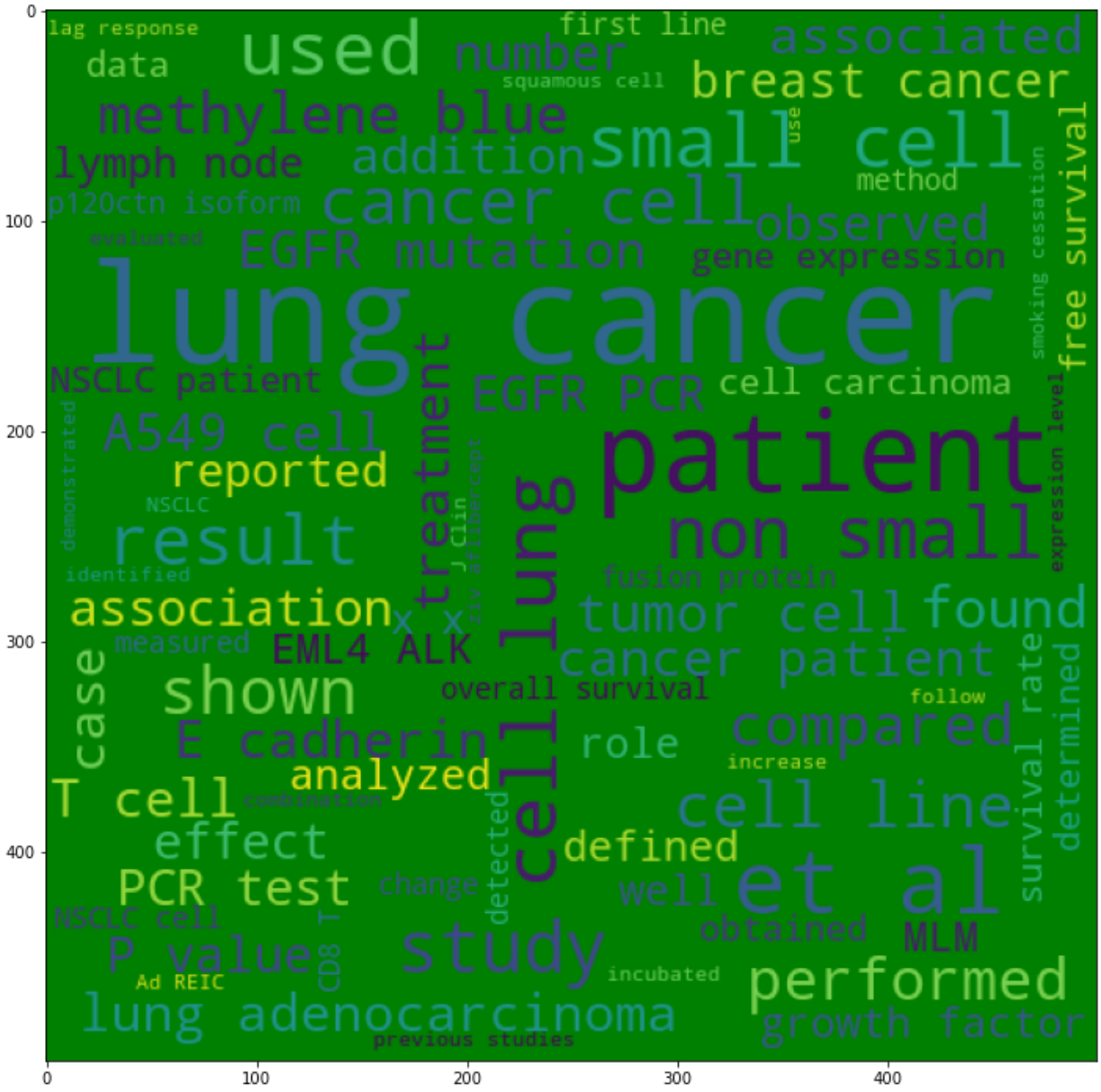
In [13]:
```
#Visualize the WordCloud in the Colon_Cancer from the dataset
plt.figure(figsize=(12,15))
wc=WordCloud(height=500,width=500,min_font_size=10,background_color='red')
wc_1=wc.generate(data[data['Label']=='Colon_Cancer']['Text'].str.cat(sep=" "))
plt.imshow(wc_1)
```

Out[13]:
```
<matplotlib.image.AxesImage at 0x7fc925a183d0>
```



In [14]:
```
#Visualize the WordCloud in the Colon_Cancer from the dataset
plt.figure(figsize=(12,15))
wc=WordCloud(height=500,width=500,min_font_size=10,background_color='green')
wc_2=wc.generate(data[data['Label']=='Lung_Cancer']['Text'].str.cat(sep=" "))
plt.imshow(wc_2)
```

Out[14]:
```
<matplotlib.image.AxesImage at 0x7fc919dd5fd0>
```



DATA PREPROCESSING

Remove the numbers in the dataset

In [15]:
```
#Define the function for remove the numbers in the dataset
def remove_number(text):
    #Create loop for remove the digits in the dataset
    text="".join([i for i in text if not i.isdigit()])
    #Return to the dataset
    return text
#Apply the above function to the dataset
data['Text']=data['Text'].apply(remove_number)
```

Clean the text

In [16]:
```
#Define the function to clean the text
def clean_text(text):
    pattern = r'[^a-zA-Z\s]'
    text=re.sub(pattern,'',text)
    return text
#Apply to the function to the dataset
data['Text']=data['Text'].apply(clean_text)
```

Remove the stopwords

In [17]:
```
#Create function to the remove the stopwods
names = ['Colon_Cancer', 'Lung_Cancer', 'Thyroid_Cancer']
def clean_stop(text):
    stop_words = stopwords.words('english')
    for name in names:
        stop_words.append(name)
    return " ".join([w.lower() for w in text.split() if w.lower() not in stop_
#And finally apply the above function to the dataset
data['Text']=data['Text'].apply(clean_stop)
```

Tokenizeation

In [18]:
```
#Define the tokenize function
def tokenize(d):
    return word_tokenize(d)
data['Text']=data['Text'].apply(tokenize)
```

In [19]:
```
data.head()
```

Out[19]:

	Label	Text
0	Thyroid_Cancer	[thyroid, surgery, children, single, instituti...
1	Thyroid_Cancer	[adopted, strategy, used, prior, years, based,...
2	Thyroid_Cancer	[coronary, arterybypass, grafting, thrombosis,...
3	Thyroid_Cancer	[solitary, plasmacytoma, sp, skull, uncommon, ...
4	Thyroid_Cancer	[study, aimed, investigate, serum, matrix, met...

Remove the special characters

In [20]:
```
#removing special character
def remove_special_char(list):
    y=[]
    for string in list:
        if string.isalnum():
            y.append(string)
    return y
data['Text']=data['Text'].apply(lambda x:remove_special_char(x))
```

In [21]:
```
data.head()
```

Out[21]:

	Label	Text
0	Thyroid_Cancer	[thyroid, surgery, children, single, instituti...
1	Thyroid_Cancer	[adopted, strategy, used, prior, years, based,...
2	Thyroid_Cancer	[coronary, arterybypass, grafting, thrombosis,...
3	Thyroid_Cancer	[solitary, plasmacytoma, sp, skull, uncommon, ...
4	Thyroid_Cancer	[study, aimed, investigate, serum, matrix, met...

Stemming process

In [22]:
```
#Create a function to the stemming processs
ps=PorterStemmer()
def stemming(list):
    #Create a empty list
    y=[]
    #Create a for loop for text in list
    for text in list:
        #Then i finallay append to the empty list
        y.append(ps.stem(text))
    #then return to the empty list
    return y
data['Text']=data['Text'].apply(lambda x:stemming(x))
#join the words
data['Text']=data['Text'].apply(lambda x:" ".join(x))
```

In [23]:
```
data.head()
```

Out[23]:

	Label	Text
0	Thyroid_Cancer	thyroid surgeri children singl institut osama ...
1	Thyroid_Cancer	adopt strategi use prior year base four exclus...
2	Thyroid_Cancer	coronari arterybypass graft thrombosi brin bri...
3	Thyroid_Cancer	solitari plasmacytoma sp skull uncommon clinic...
4	Thyroid_Cancer	studi aim investig serum matrix metalloprotein...

Modeling

In [24]:
```
#Divided the data into two variables
X=data['Text']
y=data['Label']
```

In [25]:
```
#install the TfidfVectorizer
vector=TfidfVectorizer()
#Fit the X data to the TfidfVectorizer
vector.fit(X)
```

```
#And transform the
X=vector.transform(X)
```

In [26]:
```
#And Divided the data into traing and testing and finally split the data and
X_train,X_test,y_train,y_test=train_test_split(X,y,test_size=0.25,random_state
```

LogisticRegression

In [27]:
```
#Install the logisticregression model
logistic=LogisticRegression()
#And fit the model to the train data
logistic.fit(X_train,y_train)
```

Out[27]:
```
LogisticRegression()
```

In [28]:
```
#Prediction of the logisticregression algorithm
logistic_pred=logistic.predict(X_test)
logistic_pred
```

Out[28]:
```
array(['Colon_Cancer', 'Colon_Cancer', 'Colon_Cancer', ..., 'Lung_Cancer',
       'Thyroid_Cancer', 'Colon_Cancer'], dtype=object)
```

In [29]:
```
from sklearn.metrics import classification_report,accuracy_score,confusion_mat
```

In [30]:
```
#Check the test score and train score to the logisticregression algorithm
print(f'The Test_accuracy: {logistic.score(X_test,y_test)*100:.2f}')
#Train score for the data
print(f'The Train_accuracy: {logistic.score(X_train,y_train)*100:.2f}')
#Check the accuracy_score to the model
print(f'The Accuracy_score: {accuracy_score(y_test,logistic_pred)*100:.2f}')
```

```
The Test_accuracy: 93.29
The Train_accuracy: 95.90
The Accuracy_score: 93.29
```

Classification_report and Confusion_matrix

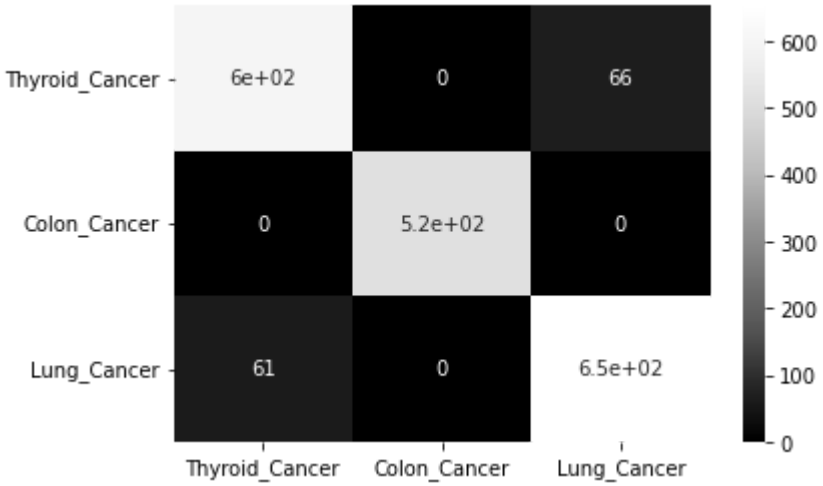
In [31]:
```
#Classification report
print(classification_report(y_test,logistic_pred))
#confusion_matrix
cn=confusion_matrix(y_test,logistic_pred)
sns.heatmap(cn,annot=True,cmap='Greys_r',xticklabels=['Thyroid_Cancer','Colon_
```

```
                precision    recall  f1-score   support

  Colon_Cancer       0.91      0.90      0.90       664
   Lung_Cancer       1.00      1.00      1.00       515
Thyroid_Cancer       0.91      0.91      0.91       714

      accuracy                           0.93      1893
     macro avg       0.94      0.94      0.94      1893
  weighted avg       0.93      0.93      0.93      1893
```

Out[31]:
```
<AxesSubplot:>
```



DecisionTreeClassifier

In [32]:
```
#Import the DecisionTreeClassifier algortihm
from sklearn.tree import DecisionTreeClassifier
#install the DecisionTreeClassifier model
tree=DecisionTreeClassifier()
#Fit the train data to the model
tree.fit(X_train,y_train)
```

Out[32]:
```
DecisionTreeClassifier()
```

In [33]:
```
#Prediction of the DecisionTreeClassifier algorithm
tree_pred=tree.predict(X_test)
tree_pred
```

Out[33]:
```
array(['Colon_Cancer', 'Colon_Cancer', 'Thyroid_Cancer', ...,
       'Lung_Cancer', 'Thyroid_Cancer', 'Colon_Cancer'], dtype=object)
```

In [34]:
```
#Check the test score and train score to the DecisionTreeClassifier algorithm
print(f'The Test_accuracy: {tree.score(X_test,y_test)*100:.2f}')
#Train score for the data
print(f'The Train_accuracy: {tree.score(X_train,y_train)*100:.2f}')
#Check the accuracy_score to the model
print(f'The Accuracy_score: {accuracy_score(y_test,tree_pred)*100:.2f}')
```

```
The Test_accuracy: 100.00
The Train_accuracy: 100.00
The Accuracy_score: 100.00
```

Classification_report and Confusion_matrix

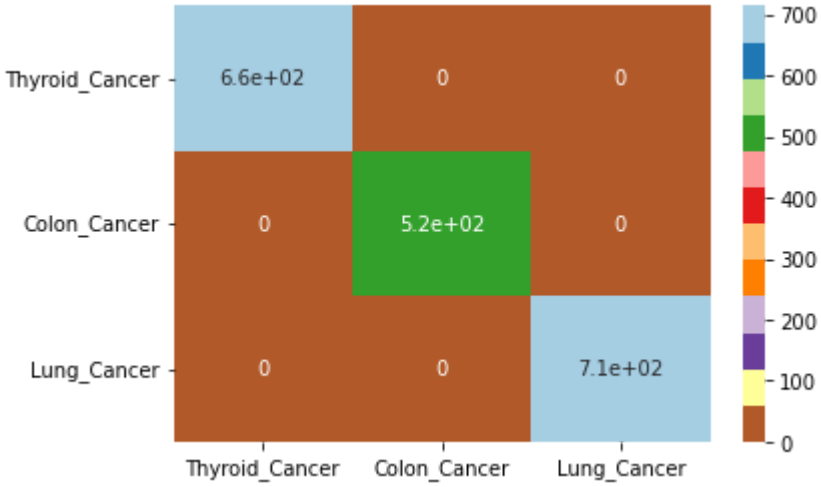
In [35]:
```
#Classification report
print(classification_report(y_test,tree_pred))
#confusion_matrix
cn=confusion_matrix(y_test,tree_pred)
sns.heatmap(cn,annot=True,cmap='Paired_r',xticklabels=['Thyroid_Cancer','Colon
```

```
                precision    recall  f1-score   support

  Colon_Cancer       1.00      1.00      1.00       664
   Lung_Cancer       1.00      1.00      1.00       515
Thyroid_Cancer       1.00      1.00      1.00       714

      accuracy                           1.00      1893
     macro avg       1.00      1.00      1.00      1893
  weighted avg       1.00      1.00      1.00      1893
```

Out[35]:
```
<AxesSubplot:>
```



RandomForestClassifier

In [36]:
```
#Import the RandomForestClassifier algortihm
from sklearn.ensemble import RandomForestClassifier
#install the DecisionTreeClassifier model
random=RandomForestClassifier()
#Fit the train data to the model
random.fit(X_train,y_train)
```

Out[36]:
```
RandomForestClassifier()
```

In [37]:
```
#Prediction of the RandomForestClassifier algorithm
random_pred=random.predict(X_test)
random_pred
```

Out[37]:
```
array(['Colon_Cancer', 'Colon_Cancer', 'Thyroid_Cancer', ...,
       'Lung_Cancer', 'Thyroid_Cancer', 'Colon_Cancer'], dtype=object)
```

In [38]:
```
#Check the test score and train score to the RandomForestClassifier algorithm
print(f'The Test_accuracy: {random.score(X_test,y_test)*100:.2f}')
#Train score for the data
print(f'The Train_accuracy: {random.score(X_train,y_train)*100:.2f}')
#Check the accuracy_score to the model
print(f'The Accuracy_score: {accuracy_score(y_test,random_pred)*100:.2f}')
```

```
The Test_accuracy: 100.00
The Train_accuracy: 100.00
The Accuracy_score: 100.00
```

Classification_report and Confusion_matrix

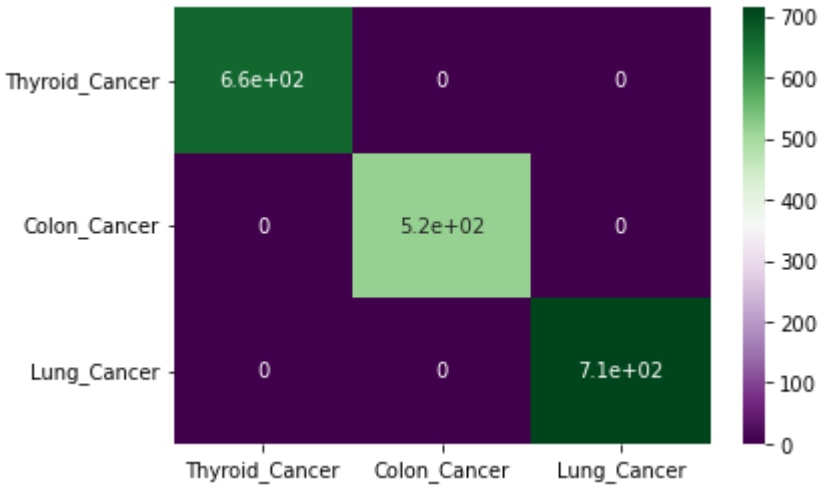
In [39]:
```
#Classification report
print(classification_report(y_test,random_pred))
#confusion_matrix
cn=confusion_matrix(y_test,random_pred)
sns.heatmap(cn,annot=True,cmap='PRGn',xticklabels=['Thyroid_Cancer','Colon_Can
```

```
                precision    recall  f1-score   support

  Colon_Cancer       1.00      1.00      1.00       664
   Lung_Cancer       1.00      1.00      1.00       515
Thyroid_Cancer       1.00      1.00      1.00       714

      accuracy                           1.00      1893
     macro avg       1.00      1.00      1.00      1893
  weighted avg       1.00      1.00      1.00      1893
```

Out[39]:
```
<AxesSubplot:>
```



MultinomialNB

In [40]:
```
#Import the MultinomialNB algorithm to train the our model
from sklearn.naive_bayes import MultinomialNB
#install the model
multinomial=MultinomialNB()
#fit the train data to our model
multinomial.fit(X_train,y_train)
```

Out[40]:
```
MultinomialNB()
```

In [41]:
```
#Prediction to the test data MultinomialNB
multinomial_pred=multinomial.predict(X_test)
multinomial_pred
```

Out[41]:
```
array(['Colon_Cancer', 'Colon_Cancer', 'Thyroid_Cancer', ...,
       'Lung_Cancer', 'Thyroid_Cancer', 'Colon_Cancer'], dtype='<U14')
```

In [42]:
```
#Check the test score and train score to the MultinomialNB algorithm
print(f'The Test_accuracy: {multinomial.score(X_test,y_test)*100:.2f}')
```

```
#Train score for the data
print(f'The Train_accuracy: {multinomial.score(X_train,y_train)*100:.2f}')
#Check the accuracy_score to the model
print(f'The Accuracy_score: {accuracy_score(y_test,multinomial_pred)*100:.2f}'
```

```
The Test_accuracy: 90.86
The Train_accuracy: 93.62
The Accuracy_score: 90.86
```

Classification_report and Confusion_matix

In [43]:

```
#Classification report
print(classification_report(y_test,multinomial_pred))
#confusion_matrix
cn=confusion_matrix(y_test,multinomial_pred)
sns.heatmap(cn,annot=True,cmap='Pastel1_r',xticklabels=['Thyroid_Cancer','Colo
```

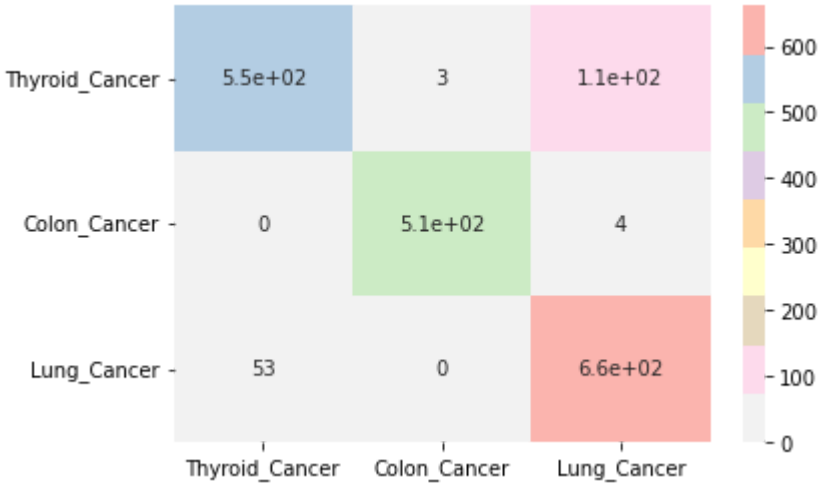
```
                precision    recall  f1-score   support

  Colon_Cancer       0.91      0.83      0.87       664
   Lung_Cancer       0.99      0.99      0.99       515
Thyroid_Cancer       0.85      0.93      0.89       714

      accuracy                           0.91      1893
     macro avg       0.92      0.91      0.92      1893
  weighted avg       0.91      0.91      0.91      1893
```

Out[43]:

```
<AxesSubplot:>
```



XGBClassifier

In [44]:

```
#Import the XGBClassifier model and install the model
from xgboost import XGBClassifier
#install the XGBClassifier
xgb=XGBClassifier()
#And finally fit the data to train data
xgb.fit(X_train,y_train)
```

Out[44]:
```
XGBClassifier(base_score=0.5, booster=None, colsample_bylevel=1,
              colsample_bynode=1, colsample_bytree=1, gamma=0, gpu_id=-1,
              importance_type='gain', interaction_constraints=None,
              learning_rate=0.300000012, max_delta_step=0, max_depth=6,
              min_child_weight=1, missing=nan, monotone_constraints=None,
              n_estimators=100, n_jobs=0, num_parallel_tree=1,
              objective='multi:softprob', random_state=0, reg_alpha=0,
              reg_lambda=1, scale_pos_weight=None, subsample=1,
              tree_method=None, validate_parameters=False, verbosity=None)
```

In [45]:
```
#Prediction to the test data XGBClassifier
xgb_pred=xgb.predict(X_test)
xgb_pred
```

Out[45]:
```
array(['Colon_Cancer', 'Colon_Cancer', 'Thyroid_Cancer', ...,
       'Lung_Cancer', 'Thyroid_Cancer', 'Colon_Cancer'], dtype=object)
```

In [46]:
```
#Check the test score and train score to the XGBClassifier algorithm
print(f'The Test_accuracy: {xgb.score(X_test,y_test)*100:.2f}')
#Train score for the data
print(f'The Train_accuracy: {xgb.score(X_train,y_train)*100:.2f}')
#Check the accuracy_score to the model
print(f'The Accuracy_score: {accuracy_score(y_test,xgb_pred)*100:.2f}')
```

```
The Test_accuracy: 100.00
The Train_accuracy: 100.00
The Accuracy_score: 100.00
```

Classification_report and Confusion_matrix

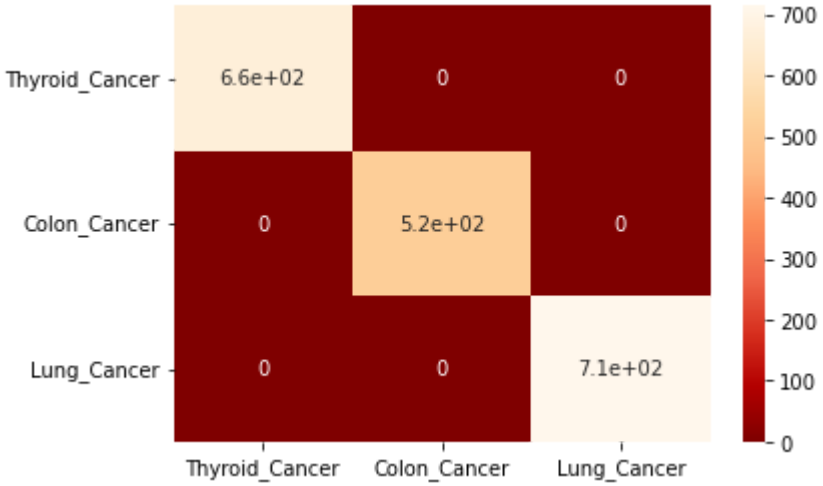
In [47]:
```
#Classification report
print(classification_report(y_test,xgb_pred))
#confusion_matrix
cn=confusion_matrix(y_test,xgb_pred)
sns.heatmap(cn,annot=True,cmap='OrRd_r',xticklabels=['Thyroid_Cancer','Colon_C
```

```
                precision    recall  f1-score   support

  Colon_Cancer       1.00      1.00      1.00       664
   Lung_Cancer       1.00      1.00      1.00       515
Thyroid_Cancer       1.00      1.00      1.00       714

      accuracy                           1.00      1893
     macro avg       1.00      1.00      1.00      1893
  weighted avg       1.00      1.00      1.00      1893
```

Out[47]:
```
<AxesSubplot:>
```



Deep Learning models

In [48]:
```
#Import the necessary librairys
from tensorflow.keras.utils import to_categorical
from gensim.models import Word2Vec
from gensim.models.keyedvectors import KeyedVectors
import time
from keras.layers import Dense, Input, Flatten, LSTM, Bidirectional,Embedding,
from keras.layers import Conv1D, MaxPooling1D, Embedding
from keras.models import Sequential, load_model
from keras import losses
from tensorflow.keras.optimizers import Adam
from tensorflow.keras.models import Model
from keras.utils import pad_sequences
from keras.utils.vis_utils import plot_model
from keras.callbacks import EarlyStopping
from keras.preprocessing.text import Tokenizer
```

In [61]:
```
#Create the dummies values to the dataset and divided the data train test spli
y = pd.get_dummies(data.Label)
X_trn, X_tst, y_trn, y_tst = train_test_split(X, y, test_size=0.2, random_stat
X_trn, X_vld, y_trn, y_vld = train_test_split(X_trn, y_trn, test_size=0.3, ran
```

In [50]:
```
#Creat the maX_word and max_len variables
max_words = 5000
max_len = 300
#Then create a function for the panding the text data
def tokenize_pad_sequences(text):
    # Text tokenization
    tokenizer = Tokenizer(num_words=max_words, lower=True, split=' ')
    tokenizer.fit_on_texts(text)
    # Transforms text to a sequence of integers
    X = tokenizer.texts_to_sequences(text)
    # Pad sequences to the same length
    X = pad_sequences(X, padding='post', maxlen=max_len)
    # return sequences
    return X, tokenizer

print('Before Tokenization & Padding \n', data['Text'][0],'\n')
```

```
X, tokenizer = tokenize_pad_sequences(data['Text'])
print('After Tokenization & Padding \n', X[0])
```

Before Tokenization & Padding

thyroid surgeri children singl institut osama ibrahim almosallama ali aseerib ahm alhumaida ali alzahran saif alsobhib saud alshanafeybfrom adepart surgeri colleg medicin qassim univers buraidah al qassim saudi arabia bdepart surgeri king faisal specialist hospit research center riyadh saudi arabia cdepart medi cin king faisal specialist hospit research center riyadh saudi arabia correspo nd dr osama ibrahim almosallam depart surgeri colleg medicin qassim univers po box buraidah al qassim saudi arabia osamaiaahotmailcom orcid orcid citat almos allam oi aseeri alhumaid alzahrani alsobhi alshanafey thyroid surgeri children singl institut ann saudi med receiv januari accept may publish august copyrigh t copyright annal saudi medicin saudi arabia access creativ common attribution noncommercialnoderiv intern licens cc byncnd detail access httpcreativecommon licensesbyncndfund nonebackground data thyroid surgeri children scarceobject a nalyz outcom data thyroid surgeri pediatr populationdesign medic record review set tertiari health care institutionpati method collect demograph clinic data patient year younger thyroid surgeri period descript data presentedmain outcom measur indic thyroidectomi thyroid patholog complic length stay radioact iodin treatment recurrencessampl size result patient underw thyroidectomi procedur f emal mean age oper year associ multipl endocrin neoplasia type histori radiat exposur eightyon patient fine needl aspir fna correl final histopatholog case sixtysix patient malign cancer papillari patient neck dissect lymph node metas tasi distant metastas lung procedur includ total thyroidectomi hemithyroidecto mi complet subtot thyroidectomi twentythre patient develop hypocalcemia perman unilater recurr laryng nerv injuri perman patient follow mean durat month medi an month patient thyroid cancer receiv radioact iodin recurr malign commonest indic thyroid surgeri children fna highli diagnost hypocalcemia recurr laryng nerv injuri signific complic recurr rate thyroid cancer limit retrospectivecon flict interest noneorigin ann saudi med julyaugust wwwannsaudimednet cthyroid diseas requir surgeri rel uncommon children compar adult preval palpabl thyroi d nodul children rang sporad welldifferenti thyroid cancer common endocrin mal ign children account pediatr cancer prepubert age group cancer adolesc age yea r common indic thyroid surgeri children vari among publish studi thyroidectomi malign condit rise data children throughout world rel scarc object studi analy z clinic data outcom thyroid surgeri larg seri children treat singl center kin g faisal specialist hospit research center kfshrc riyadhpati method approv ins titut review board irb kfshrc medic record patient year old younger underw thy roid surgeri retrospect review elect includ patient year ensur reason followup period patient studi identifi search oper room log procedur involv thyroid gla nd specifi age groupdemograph data clinic featur surgic outcom collect specif data obtain includ age oper gender famili histori present symptom histori radi at exposur presenc multipl endocrin neoplasia type men thyroid function test p resenc size thyroid nodul ultrasound presenc lymph node metastasi distant meta stasi fine needl aspir fna cytolog surgic procedur final histopatholog length followup outcom analyz postop complic includ transient perman hypocalcemia tra nsient perman recurr laryng nerv paralysi wound infect hematoma length stay ra dioact iodin treatment recurr thyroid procedur seri includ hemithyroidectomi s ubtot total complet thyroidectomi surgeri perform either endocrin adult surgeo n pediatr surgeon intraop nerv monitor use earli seri procedur perform adult e ndocrin surgeon late combin approach adopt pediatr surgeon adult endocrin surg eon collabor case proceduresth normal rang laboratori regardless symptom trans ient hypocalcemia identifi last less month perman hypocalcemia consid serum ca lcium level remain normal rang patient continu calcium supplement month surger i patient famili histori men underw genet test ret protooncogen confirm diagno si patient underw complet thyroidectomi preoper postop vocal cord assess otola ryngolog clinic descript data gener comparison conduct use test continu data c hisquar fisher exact test proportionsresultsbetween patient underw surgic proc edur patient underw two procedur thyroid diseas institut eighti patient femal mean age oper year median year rang year common indic thyroidectomi thyroid no dul present case tabl mean sd size thyroid nodul mm case associ men syndrom fi nal patholog two patient men syndrom show medullari thyroid cancer mtc remain

patient prophylact procedur develop mtc none patient histori radiat exposur ei ghtyon patient fna correl final histopatholog case three case toxic adenoma on e case grave diseas requir fna remain case underw fna anoth institut fna repea t institut came complet thyroidectomi document patholog malign first surgeri a noth hospitalth common diagnos includ papillari thyroid cancer multinodular go iter colloid tabl indic thyroidectomi patient indicationnodulen men prophylaxi shyperthyroidismmultinodular goitercomplet thyroidectomi hypocalcemia defin ca lcium level data number origin pediatr thyroid surgeryann saudi med julyaugust wwwannsaudimednet cnodul tabl surgic procedur includ total thyroidectomi hemit hyroidectomi complet thyroidectomi subtot thyroidectomi neck dissect perform p atient oper complic observ patient common complic hypocalcemia transient perma n tabl thyroid patholog patientspathologyn benignnorm thyroid tissuecolloid no dulecystadenomathyroiditisgrav diseasethyroid cancerpapillaryfollicularmedulla ryhurthleanaplastictotaldata number tabl benign malign lesion patientsbenignnm alignantn valu age meanyearsgend malefemalepres nodulehypocalcemiarecurr laryn g nerv palsybleedinghematomawound infectiontrach injuryoveral complicationsmea n length stay daysmen recurr laryng nerv palsi transient perman unilater tabl patient malign lesion lymph node metastasi patient distant metastas lung none patient develop postop bleed wound infect tracheal injuri patient follow mean month median rang month radioact iodin treatment deliv patient malign lesion p atient recurr local recurr local distant recurr lung three case receiv radioac t iodin rai recurr one case low risk recurr receiv rai recurr one case medulla ri thyroid cancer receiv rai remain five case clear data whether patient recei v rai recurr local recurr underw resect except one patient lost follow mortal studi discussionth common indic thyroidectomi seri thyroid nodul correl previo us publish report pediatr popul children thyroid nodul estim fourfold higher r isk develop thyroid cancer compar adult high incid malign seri suggest childre n thyroid nodul care evaluatedfna valuablemethod preoper evalu thyroid nodul h owev limit routin use fna children includ need sedat sampl error limit avail e xperienc cytopathologist mani previou studi report high sensit specif fna eval u thyroid nodul children correl findingsour data show lymph node metastasi thy roid cancer case support notion children thyroid cancer frequent present exten s diseas adult lymphnod involv diagnosi seen children compar adult differenti thyroid cancer hospit largest referr center saudi arabia especi oncolog case m ay explain larg number lymph node distant metastasi cohortth common complic re port thyroidectomi children hypoparathyroid incid rang origin pediatr thyroid surgeryann saudi med julyaugust wwwannsaudimednet ccorrespond result report hy pocalcemia tabl one studi found total thyroidectomi central bilater neck disse ct grave diseas malign risk factor hypocalcemia thyroid surgeri cohort postop hypocalcemia note malign case fail reach statist signific moreov signific diff er benign malign case term mean age gender distribut recurr laryng nerv injuri overal complic find report previous multipl studi recent year found invers rel ationship surgeon volum complic rate similar data pediatr popul lack one studi found highvolum endocrin surgeon better outcom shorter length stay lower cost thyroidectomi parathyroidectomi children compar pediatr surgeon gener surgeon otolaryngologist scheumann colleagu also conclud collabor approach pediatr end ocrin surgeon would better outcom led author suggest combin approach endocrin pediatr surgeon addit pediatr endocrinologist may optim care children surgic t hyroid diseas given low number pediatr patient data allow comparison differ ap proach given late adopt combin approach recurr rate thyroid cancer children th yroidectomi vari wide report studi rang cohort studi explor predictor recurr l ymph node involv multipl nodul male gender younger age histolog subtyp advanc tumor stage risk factor associ recurr studi patient malign lesion receiv rai a lthough conflict data regard indic postop rai treatment lowrisk patient curren t recommend lowrisk patient treat without raither limit studi retrospect natur may affect valid qualiti data small number case categori enabl us compar group explor predictor rel factor hand studi add scarc data thyroid surgeri pediatr age group malign commonest indic thyroid surgeri children fna highli diagnost hypocalcemia recurr laryng nerv injuri signific complic cancerrel death extrem rare recurr uncommon signific number patient malign case receiv rai treatmento

rigin pediatr thyroid surgeryann saudi med julyaugust wwwannsaudimednet crefer trowbridg fl matovinov mclaren gd nichaman mz iodin goiter children pediatr ri e lag melbert krapcho stinchcomb dg howlad horner mj et al seer cancer statist review bethesda nation cancer institut base novemb seer data submiss chen yh m asiako pt gaz rd hodin ra parangi randolph gw et al pediatr thyroidectomi high volum thyroid surgeri center risk factor postop hypocalcemia pediatr surg aug wood jh partrick da barham hp bensard dd traver hs bruni jl et al pediatr thyr oidectomi collabor surgic approach pediatr surg may scholz smith jr chaignaud shamberg rc huang sa thyroid surgeri children hospit boston year singleinstitu t experi pediatr surg mar josefson zimmerman thyroid nodul cancer children ped iatr endocrinol rev sep hame zacharin mr chang face paediatr adolesc thyroid c ancer paediatr child health lugovicent ortiz vn irizarri camp ji pagn pediatr thyroid nodul manag era fine needl aspirationj pediatr surg mussa de andrea mo tta mormil palestini corria predictor malign children thyroid nodul pediatr oc t amirazodi propst ej chung ct parra da wasserman jd pediatr thyroid fna biops i outcom impact manag year tertiari care center cancer cytopathol partyka kl h uang ec cramer hm chen wu hh histolog clinic followup thyroid fineneedl aspir pediatr patient cancer cytopathol sinha ck decoppi pierro brain hindmarsh butl er et al thyroid surgeri children clinic outcom eur pediatr surg oct kundel th ompson gb richard ml qiu lx cai schwenk fw et al pediatr endocrin surgeri year experi mayo clinic clin endocrinol metab februari jiang newburi ro newfield rs pediatr thyroid surgeri manag thyroid nodulesan institut experi featur year pe riod int pediatr endocrinol burk jf sippel rs chen evolut pediatr thyroid surg eri tertiari medic center surg re alqahtani kh tunio al asiri aljohani nj bayo umi riaz et al clinicopatholog treatment outcom differenti thyroid cancer saud i children adult otolaryngol head neck surg nov kluijfhout wp van beek dj verr ijn stuart aa lodewijk valk gd van der zee dc et al postop complic prophylact thyroidectomi young patient multipl endocrin neoplasia type medicin baltimor r aval mv brown chin ac zimmerman angelo reynold total thyroidectomi benign dise as pediatr patientfeas safe pediatr surg stavraki ai ituart ph ko cy yeh mw su rgeon volum predictor outcom inpati outpati endocrin surgeri surgeri sosa ja b owman hm tielsch jm pow nr gordon ta udelsman import surgeon experi clinic eco nom outcom thyroidectomi ann surg tuggl ct roman sa wang ts boudouraki thoma u delsman et al pediatr endocrin surgeri oper children surgeri dec park jeong js ryu hr lee park jh kang et al differenti thyroid carcinoma children adolescent syear experi yonsei univers health system korean med sci palmer ba zarroug ae poley rn kollar jp moir cr papillari thyroid carcinoma children risk factor co mplic diseas recurr pediatr surg wada sugino mimura nagahama kitagawa shibuya et al pediatr differenti thyroid carcinoma stage risk factor analysi diseas fr ee surviv bmc cancer danes gardini farsetti sciacchitano andreoli pontecorvi t hyroid carcinoma children adolesc eur pediatr astl chovanec luke katra dvorako va vlcek et al thyroid carcinoma surgeri children adolesc year experi surgeri pediatr thyroid lymph node metastas carcinoma int pediatr otorhinolaryngol cha ukar da rangarajan nair nadkarni ms pai ps dcruz ak et al pediatr thyroid canc er surg oncol dzodic buta markov gavrilo matov milovanov et al surgic manag we lldifferenti thyroid carcinoma children adolesc year experi singl institut ser bia endocr scheumann gf gimm wegen hundeshagen drall prognost signific surgic manag locoregion papillari thyroid cancer world surg shi rl qu yang sw tumor s ize interpret predict cervic lymph node metastasi use differenti thyroid cance r risk model onco target ther zimmerman hay id gough ir goellner jr ryan jj gr ant cs et al papillari thyroid carcinoma children adult longterm followup pati ent conserv treat one institut three decad surgeri collini mattavelli pellegri nelli barisella ferrari massimino papillari carcinoma thyroid gland childhood adolesc morpholog subtyp biolog behavior prognosi clinicopatholog studi sporad case treat singl institut year period surg pathol borsonchazot causeret lifant jc augro berger peix jl predict factor recurr seri children adolesc differenti thyroid cancer world surg baumgarten hd bauer aj isaza mostoufimoab kazahaya a dzick ns surgic manag pediatr thyroid diseas complic rate thyroidectomi childr en hospit philadelphia highvolum pediatr thyroid center journal pediatr surger i oct kurzawinski tr de coppi thyroidectomi children inpediatr surgeri pp spri

```
nger berlin heidelberg franci waguespack sg bauer aj angelog benvenga et al ma
nag guidelin children thyroid nodul differenti thyroid cancer american thyroid
associ guidelin task forc pediatr thyroid cancer thyroid volum number origin p
ediatr thyroid surgeryann saudi med julyaugust wwwannsaudimednet

After Tokenization & Padding
 [ 233 2100 2960   73  189 2762  673  124 3270 1474   20 1209 1688 1209
  737 2585 1315 1663 3469 1343  285 1322  195 2045 2702 2100  134  134
 1949 3868 1477 3517 3759  157 1343  186   28 1929  195 3270 1250  737
  331  888  356 2849 3625   31   89 1209 2100  134  748 1236  134 4836
 2099 2392  462 1053 2099 2002 3781   31   89  159  176   92 1236  186
  114  112   99 3172  355  758 1900  785 2575  972 1360  176   92 1236
   43   35  498   20  365 1209  737   31   89 1209  159  176   92  106
   43   35   18   20  799   47  291    2  176   92 1236 3218 1412 1209
   31   89  176   92  134 1236 3218   61  186  134 1209  176  352  257
  593   92  574 1209 1530  875 1074 2796   31   89 1209  176    2  737
  492   31   89  350  413  176   92 1236 3218   61  186  446  326 4677
  298   53  350  413 3869 1360  176    2  740  737 2414 2183 4808  821
 1575    8  212  774  120  495  352  257  162    4  159  176    2   43
   42   54 2556 1835 1099 1790 2542  999 1985   31   89 1360  176   92
 1236  604  943  647    3 1371  110   39  326  138 1776  134 1360   92
  176 1279 3626 3218 1156  576  204  933  243 1142    5 4591   34  110
  446  326   61  404  737 1613 2234 2003  120   35  365  983 1236 3218
  159  176    2  740  737 2880 2175 2037  350  413 1209  176   20  498
   60 3270 1236  128 1209  176  338  451 1209  134 3579 1527  526 3270
 1236  134  584 3543 4171 4809 2175   31   89  413  566 1236  176  613
  159  176    2  804  176   19  566 2430 2653 1209  176    2  176  285
   58  327 1209  176 2795  355]
```

In [51]:
```
#Vocab_size
vocab_size = 5000
#Embeddig_size
embedding_size = 32
epochs=50
#Let's install the Sequential model
model= Sequential()
#then add the Embedding to the model with the vocab_size and input_length
model.add(Embedding(vocab_size, embedding_size, input_length=max_len))
#Add the convlution 1d to the model with filters 32 and padding, relu activati
model.add(Conv1D(filters=32, kernel_size=3, padding='same', activation='relu')
#Add another macpooling layer to the model
model.add(MaxPooling1D(pool_size=2))
#Add the LSTM to the model
model.add(Bidirectional(LSTM(32)))
#Add the dropout function to the model
model.add(Dropout(0.4))
#And finally add to the dense layer to the model
model.add(Dense(3, activation='sigmoid'))

plot_model(model, show_shapes = True)
```

```
2022-09-10 21:17:07.968023: I tensorflow/core/platform/cpu_feature_guard.cc:19
3] This TensorFlow binary is optimized with oneAPI Deep Neural Network Library
(oneDNN) to use the following CPU instructions in performance-critical operati
ons:  AVX2 AVX512F AVX512_VNNI FMA
To enable them in other operations, rebuild TensorFlow with the appropriate co
mpiler flags.
You must install pydot (`pip install pydot`) and install graphviz (see instruc
tions at https://graphviz.gitlab.io/download/) for plot_model/model_to_dot to
work.
```

In [52]:
```
#Let's compile the model the categorical_crossentropy loss function and adam o
model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['acc
print(model.summary())
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 embedding (Embedding)       (None, 300, 32)           160000

 conv1d (Conv1D)             (None, 300, 32)           3104

 max_pooling1d (MaxPooling1D  (None, 150, 32)          0
 )

 bidirectional (Bidirectiona  (None, 64)               16640
 l)

 dropout (Dropout)           (None, 64)                0

 dense (Dense)               (None, 3)                 195

=================================================================
Total params: 179,939
Trainable params: 179,939
Non-trainable params: 0
_________________________________________________________________
None
```

In [62]:
```
#Create a earlystopping function
es = EarlyStopping(monitor = 'val_loss', patience=5)
batch_size = 64
#And fit the model to thre train and validation data
history = model.fit(X_trn, y_trn,
                    validation_data=(X_vld, y_vld),
                    batch_size=batch_size, epochs=epochs, verbose=1,
                    callbacks = [es])
```

```
Epoch 1/50
67/67 [==============================] - 8s 74ms/step - loss: 0.9791 - accurac
y: 0.4713 - val_loss: 0.6943 - val_accuracy: 0.6247
Epoch 2/50
67/67 [==============================] - 4s 67ms/step - loss: 0.4512 - accurac
y: 0.8351 - val_loss: 0.2462 - val_accuracy: 0.9290
Epoch 3/50
67/67 [==============================] - 4s 65ms/step - loss: 0.1701 - accurac
y: 0.9566 - val_loss: 0.1000 - val_accuracy: 0.9719
Epoch 4/50
67/67 [==============================] - 4s 65ms/step - loss: 0.1443 - accurac
y: 0.9533 - val_loss: 0.1154 - val_accuracy: 0.9565
Epoch 5/50
67/67 [==============================] - 4s 65ms/step - loss: 0.0598 - accurac
y: 0.9866 - val_loss: 0.0546 - val_accuracy: 0.9868
Epoch 6/50
67/67 [==============================] - 4s 67ms/step - loss: 0.0452 - accurac
y: 0.9882 - val_loss: 0.0473 - val_accuracy: 0.9840
Epoch 7/50
67/67 [==============================] - 4s 65ms/step - loss: 0.0316 - accurac
y: 0.9925 - val_loss: 0.0397 - val_accuracy: 0.9884
Epoch 8/50
67/67 [==============================] - 4s 65ms/step - loss: 0.0411 - accurac
y: 0.9917 - val_loss: 0.0413 - val_accuracy: 0.9890
Epoch 9/50
67/67 [==============================] - 5s 68ms/step - loss: 0.0381 - accurac
y: 0.9913 - val_loss: 0.0395 - val_accuracy: 0.9890
Epoch 10/50
67/67 [==============================] - 4s 66ms/step - loss: 0.0298 - accurac
y: 0.9913 - val_loss: 0.0315 - val_accuracy: 0.9917
Epoch 11/50
67/67 [==============================] - 4s 66ms/step - loss: 0.0215 - accurac
y: 0.9932 - val_loss: 0.0360 - val_accuracy: 0.9917
Epoch 12/50
67/67 [==============================] - 4s 66ms/step - loss: 0.0239 - accurac
y: 0.9939 - val_loss: 0.0372 - val_accuracy: 0.9906
Epoch 13/50
67/67 [==============================] - 5s 68ms/step - loss: 0.0251 - accurac
y: 0.9925 - val_loss: 0.0284 - val_accuracy: 0.9923
Epoch 14/50
67/67 [==============================] - 4s 66ms/step - loss: 0.0164 - accurac
y: 0.9936 - val_loss: 0.0315 - val_accuracy: 0.9912
Epoch 15/50
67/67 [==============================] - 4s 66ms/step - loss: 0.0131 - accurac
y: 0.9965 - val_loss: 0.0311 - val_accuracy: 0.9917
Epoch 16/50
67/67 [==============================] - 5s 70ms/step - loss: 0.0158 - accurac
y: 0.9943 - val_loss: 0.0264 - val_accuracy: 0.9934
Epoch 17/50
67/67 [==============================] - 5s 70ms/step - loss: 0.0136 - accurac
y: 0.9960 - val_loss: 0.0268 - val_accuracy: 0.9923
Epoch 18/50
67/67 [==============================] - 4s 67ms/step - loss: 0.0121 - accurac
y: 0.9950 - val_loss: 0.0272 - val_accuracy: 0.9945
Epoch 19/50
67/67 [==============================] - 4s 67ms/step - loss: 0.0100 - accurac
y: 0.9965 - val_loss: 0.0270 - val_accuracy: 0.9945
Epoch 20/50
67/67 [==============================] - 5s 72ms/step - loss: 0.0103 - accurac
y: 0.9960 - val_loss: 0.0273 - val_accuracy: 0.9945
```

```
Epoch 21/50
67/67 [==============================] - 5s 68ms/step - loss: 0.0106 - accurac
y: 0.9955 - val_loss: 0.0267 - val_accuracy: 0.9950
```

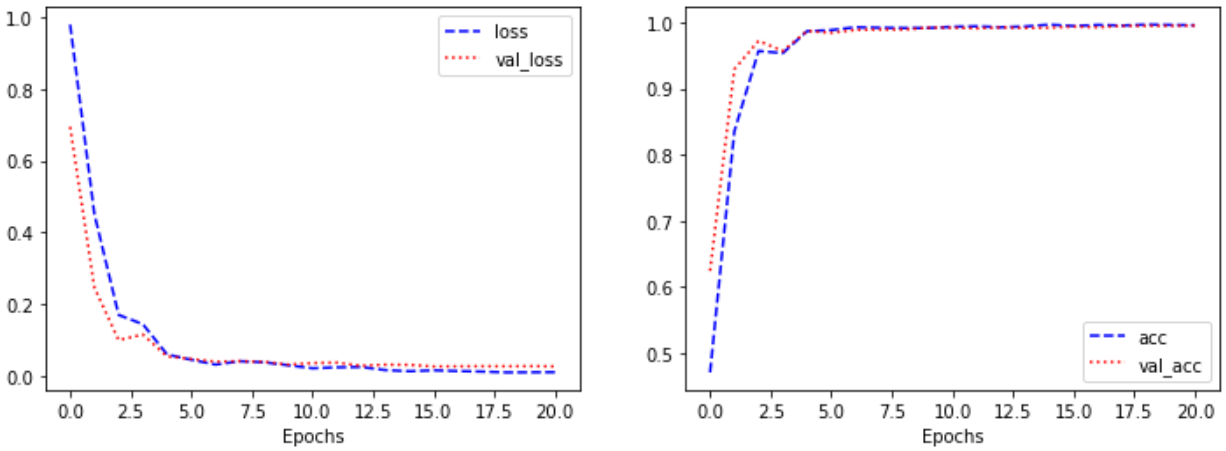
In [63]:
```
loss, accuracy = model.evaluate(X_tst, y_tst, verbose=0)

# Print metrics
print('Accuracy  : {:.4f}'.format(accuracy))
```

```
Accuracy  : 0.9941
```

In [64]:
```
#To visualize the the accuracy score to the model using the matplotlib
plt.figure(figsize=(12, 4))
#Create a subplots to the model
plt.subplot(1, 2, 1)
#plot the loss of the mode history
plt.plot(history.history['loss'], 'b--', label = 'loss')
#plot the val_loss of the history
plt.plot(history.history['val_loss'], 'r:', label = 'val_loss')
#On x-axis Epochs
plt.xlabel('Epochs')
plt.legend()
#Create a subplots to the model
plt.subplot(1, 2, 2)
#plot the accuracy of the mode history
plt.plot(history.history['accuracy'], 'b--', label = 'acc')
#plot the val_accuracy of the mode history
plt.plot(history.history['val_accuracy'], 'r:', label = 'val_acc')
#On x-axis Epochs
plt.xlabel('Epochs')
plt.legend()

plt.show()
```



Test The model predict the good result

In [65]:
```
text='the endothelium is crucial for vein integrity andprevention of thrombosi
text=[text]
text_int=vector.transform(text)
prediction=logistic.predict(text_int)
f" Biomedical text document classification is {prediction[0]}"
```

Out[65]:
```
' Biomedical text document classification is Thyroid_Cancer'
```

In [68]:
```
text='bacteroid fragili fragili produc biofilm colonis intestin tract caus ser
text=[text]
text_int=vector.transform(text)
prediction=logistic.predict(text_int)
f" Biomedical text document classification is {prediction[0]}"
```

Out[68]:
```
' Biomedical text document classification is Colon_Cancer'
```

In [67]:
```
text='lack biomark identifi aggress primari tumor subset give rise metastas im
text=[text]
text_int=vector.transform(text)
prediction=random.predict(text_int)
f" Biomedical text document classification is {prediction[0]}"
```

Out[67]:
```
' Biomedical text document classification is Lung_Cancer'
```

The LogistciRegression and DecisionTreeClassifier,RandomForestClassifer give the best result to the model.

CONCLUSION

About the data

In the data we use predict the Biomedical text document classification is wether it's Thyroid_Cancer,Lung_Cancer,Colon_Cancer based on the performed basicEDA, text preprocessing, build different models, such as LogisticRegression,DecisiontreeClassification,RandomForestClassication,XGBboostClassifier,For the above model Only All Algorithms have good accuracy score compare to the other model. After that we Run deeplearning model to the dataset. And create the Sequrentil model to the data and fit the data to the model in this model we use conv1d and several input layers used and finally we use 50 epochs to the model we get the 99% accurcy_score to the deep learning model.

In [ ]:
```
```