Table S1. Category 1. Drugs for diagnosis and treatment. Original terms for early stage of Breast Cancer.

|  |  |  |
| --- | --- | --- |
|  | Term | % in the category |
|  | Cyclophosphamide | 2,42 |
|  | Eflapegrastim | 2,18 |
|  | Technetium-99 [Tc-99] | 1,69 |
|  | Mammaglobin | 1,69 |
|  | Platycodon grandiflorum (PG) | 1,69 |
|  | Steroidal | 1,45 |
|  | Estriol vaginal gel | 1,21 |
|  | Anticoagulant(s) | 0,97 |
|  | Vernonia amygdalina Delile | 0,73 |
|  | Herceptin | 0,48 |
|  | Rivaroxaban | 0,24 |
|  | CMF 9 (cyclophosphamide, methotrexate, 5-fluorouracil) | 0,24 |

Table S2. Category 2. Bioorganic chemistry. Original terms for early stage of Breast Cancer.

|  |  |  |
| --- | --- | --- |
|  | Term | % in the category |
|  | D-dimer | 1,77 |

Table S3. Category 3. Methods and technologies of diagnosis and treatment. Original terms for early stage of Breast Cancer.

|  |  |  |
| --- | --- | --- |
|  | Term | % in the category |
|  | Neoadjuvant (NAC) | 2,82 |
|  | Hazard ratio (HR) | 2,19 |
|  | Postmastectomy radiotherapy (PMRT) | 2,09 |
|  | Intraoperative radiotherapy (IORT) | 1,95 |
|  | External beam radiotherapy (EBRT) | 1,51 |
|  | Toxicity | 1,46 |
|  | Axillary lymph node dissection (ALND) | 1,41 |
|  | Cox proportional hazard | 1,22 |
|  | Multivariate analysis | 1,17 |
|  | Preoperative | 1,17 |
|  | Image | 1,12 |
|  | Endocrine therapy | 1,02 |
|  | Partial breast radiotherapy (PBI) | 1,02 |
|  | GEP test | 0,92 |
|  | TARGIT targeted intraoperative radiotherapy (TARGIT-IORT) | 0,88 |
|  | RCT (randomisert controlled trial) | 0,83 |
|  | NACT (neoadjuvant chemotherapy) | 0,83 |
|  | Cardiac | 0,73 |
|  | Breast-contour preservation (BCP) | 0,58 |
|  | Primary breast-conserving treatment (BCT) | 0,58 |
|  | Odds ratio | 0,58 |
|  | Clinical breast examination (CBE) | 0,54 |
|  | Monocyte-lymphocyte ratio (MLR) | 0,54 |
|  | Cardiotoxic | 0,54 |
|  | Lymphovascular invasion | 0,49 |
|  | Intelligent Breast Exam [iBE] | 0,49 |
|  | Modified radical mastectomy (MRM) | 0,49 |
|  | Regression | 0,49 |
|  | Systemic therapy | 0,49 |
|  | Hormonal therapy | 0,44 |
|  | Health care | 0,44 |
|  | Contrast-enhanced (CEUS) | 0,39 |
|  | Decoction | 0,39 |
|  | Deep-inspiration breath-hold (DIBH) | 0,34 |
|  | Free-breathing (FB) tangential radiotherapy | 0,34 |
|  | Thermographic | 0,34 |
|  | Traditional Chinese medicine (TCM) | 0,34 |
|  | Hypofractionate | 0,29 |
|  | IERM intraoperative evaluation of the retroareolar margin | 0,29 |
|  | Medication | 0,29 |
|  | MRI | 0,29 |
|  | Doppler | 0,24 |
|  | Cryoablation | 0,29 |
|  | Flow cytometry | 0,24 |
|  | Gray's test | 0,24 |
|  | Mammogram screening | 0,24 |
|  | Self-examination | 0,24 |
|  | CV | 0,24 |
|  | Cardiovascular | 0,24 |
|  | Brachytherapy | 0,19 |
|  | Deep learning radiomics (DLR) | 0,19 |
|  | IMRT | 0,19 |
|  | Multiphoton microscopy (MPM) | 0,19 |
|  | Oncological treatment | 0,19 |
|  | Postmastectomy | 0,19 |
|  | Postoperative radiotherapy | 0,19 |
|  | Risk model (PRM) | 0,19 |
|  | Surface | 0,19 |
|  | Targeted therapy | 0,19 |
|  | Tomotherapy | 0,19 |
|  | Perioperative | 0,19 |
|  | Endocrine treatment | 0,15 |
|  | Axillary surgery | 0,15 |
|  | First-line treatment | 0,15 |
|  | Radioisotope | 0,15 |
|  | Whole-breast radiotherapy | 0,15 |
|  | High-dose-rate (HDR) | 0,15 |
|  | Identification rate (IR) | 0,15 |
|  | Mastectomy-free survival | 0,15 |
|  | Non-invasive biomarkers | 0,15 |
|  | Novel marker | 0,15 |
|  | Palpable | 0,15 |
|  | Postoperative | 0,15 |
|  | qRT-PCR | 0,15 |
|  | Surgical interventions | 0,15 |
|  | Dissection | 0,10 |
|  | Electrophoresis | 0,10 |
|  | Initial treatment strategy | 0,10 |
|  | Neoadjuvant endocrine therapy | 0,10 |
|  | Neo-adjuvant systemic therapy (NST) | 0,10 |
|  | Postoperative therapy | 0,10 |
|  | Pre-chemotherapy | 0,10 |
|  | Preclinical models | 0,10 |
|  | Radiation dose | 0,10 |
|  | Support Vector Machine | 0,10 |
|  | Systemic chemotherapy | 0,10 |
|  | Contralateral breast dose | 0,10 |
|  | Inflammatory markers | 0,10 |
|  | Lymphocyte count | 0,10 |
|  | Omics | 0,10 |
|  | Omission of surgery | 0,10 |
|  | Operable | 0,10 |
|  | Probabilistic Neural Network | 0,10 |
|  | Single dose | 0,10 |
|  | Frozen pathology | 0,10 |
|  | High-precision radiotherapy | 0,05 |
|  | Immunoblotting | 0,05 |
|  | Long-term survival analyses | 0,05 |
|  | Mammoplasty | 0,05 |
|  | Microarray technology | 0,05 |
|  | Nanoparticle-based protein enrichment technology | 0,05 |
|  | Oncoplastic surgery | 0,05 |
|  | Postoperative histopathology | 0,05 |
|  | Preoperative imaging. | 0,05 |
|  | Primary breast surgery | 0,05 |
|  | Regional therapy | 0,05 |
|  | Shear wave elastography | 0,05 |
|  | Single-modality therapy | 0,05 |
|  | Total breast reconstruction | 0,05 |
|  | Hemostatic biomarkers | 0,05 |
|  | Systemic markers | 0,05 |
|  | Companion biomarkers | 0,05 |

Table S4. Category 4. Genetics. Original terms for early stage of Breast Cancer.

|  |  |  |
| --- | --- | --- |
|  | Term | % in the category |
|  | Circulating tumor DNA (ctDNA) | 2,70 |
|  | cfDNA | 2,28 |
|  | OCT4 | 2,28 |
|  | Genotype | 1,85 |
|  | Cancer Genome Atlas (TCGA) | 0,85 |
|  | Gene signature (GS) | 0,85 |
|  | Oncogene | 0,85 |
|  | Genetic counseling | 0,71 |
|  | NGS next-generation sequencing | 0,71 |
|  | Cell-free nucleic acid (cfna) | 0,57 |
|  | Cell-free DNA | 0,43 |
|  | tiRNAs | 0,43 |
|  | PALB2 | 0,43 |
|  | p53 | 0,28 |
|  | cerbB2 | 0,14 |
|  | ceRNA | 0,14 |
|  | Chain-encoding rnas | 0,14 |
|  | Circulating cell-free tumor (DNA or RNA) | 0,14 |
|  | CMYC | 0,14 |
|  | ctRNA | 0,14 |
|  | Gene silencing | 0,14 |
|  | Inherited mutations | 0,14 |

Table S5. Category 5. Molecular biology. Original terms for early stage of Breast Cancer.

|  |  |  |
| --- | --- | --- |
|  | Term | % in the category |
|  | HER/2 | 7,49 |
|  | HR | 6,97 |
|  | Hormone receptor | 2,97 |
|  | Neutrophil | 2,87 |
|  | Progesterone receptor (PR) (PGR) | 2,67 |
|  | Molecular subtype | 1,74 |
|  | Aurora A (kinase) | 1,33 |
|  | POAI perioperative aromatase inhibitor | 1,33 |
|  | Adenylate cyclase-associated protein (CAP1) | 1,23 |
|  | Phosphohistone H3 (phh3) | 1,13 |
|  | TILs tumor-infiltrating lymphocyte | 1,13 |
|  | Programmed death-ligand (PD-L1) | 1,03 |
|  | Phosphoprotein (SPP1) | 0,82 |
|  | Recombinant N-terminal domain of BC2L-C lectin (rbc2lcn) | 0,82 |
|  | Cell death protein (PD-1) | 0,72 |
|  | Heparin-binding epidermal growth factor (HBEGF) | 0,72 |
|  | CaSR | 0,62 |
|  | Luteinizing hormone | 0,62 |
|  | Hsa-mir-375 | 0,51 |
|  | Peroxisome proliferator-activated receptor γ | 0,51 |
|  | Mitotic activity index (MAI) | 0,51 |
|  | Protein-protein interaction | 0,41 |
|  | Pro(-)thrombin | 0,41 |
|  | Cyclin B1 | 0,31 |
|  | Exosomes | 0,31 |
|  | Immunoglobulin | 0,31 |
|  | mTOR | 0,31 |
|  | Recombinant human granulocyte colony-stimulating factor (rhg-CSF) | 0,31 |
|  | Coagulation | 0,31 |
|  | Estrone | 0,31 |
|  | Human epidermal receptor- | 0,21 |
|  | Monoclonal | 0,21 |
|  | Fibrinogen | 0,21 |
|  | Growth factor | 0,21 |
|  | EREG 2 pathway | 0,10 |
|  | ITGA2B (integrin subunit alpha iib) | 0,10 |
|  | Low-abundance proteins | 0,10 |
|  | Oncoprotein | 0,10 |
|  | Peptidyl-Prolyl cis-trans Isomerase A (PPIA) | 0,10 |
|  | Peroxiredoxin 2 (PRDX2) | 0,10 |

Table S6. Category 6. Breast cancer. Original terms for early stage of Breast Cancer.

|  |  |  |
| --- | --- | --- |
|  | Term | % in the category |
|  | SLN(B) sentinel lymph node biopsy | 4,80 |
|  | Breast cancer patient(s) | 4,29 |
|  | Sentinel | 1,75 |
|  | ER | 1,42 |
|  | LN- lymph node | 1,38 |
|  | Whole breast irradiation (WBI) | 0,65 |
|  | Node negative | 0,55 |
|  | False | 0,44 |
|  | Breast milk (BM) | 0,36 |
|  | Regional nodal irradiation (RNI ) | 0,36 |
|  | Ct1 | 0,33 |
|  | cN0 | 0,25 |
|  | Breast cancer-specific mortality (BCSM) | 0,22 |
|  | SLNM sentinel lymph node metastasis | 0,22 |
|  | Breast cancer-specific death (BCSD) | 0,18 |
|  | Breast disease | 0,18 |
|  | Breast tumors | 0,18 |
|  | Female breast cancer (FBC) | 0,18 |
|  | pN0(i-) | 0,18 |
|  | Positive lymph nodes | 0,18 |
|  | Breast cancer cases | 0,15 |
|  | Breast cancer recurrence | 0,15 |
|  | Breast conservation | 0,15 |
|  | Lymphatic | 0,15 |
|  | Operable breast cancer | 0,15 |
|  | pN0(i+) | 0,15 |
|  | Breast cancer diagnosis | 0,11 |
|  | Breast cancer mortality | 0,11 |
|  | Breast cancer risk prediction | 0,11 |
|  | Invasive breast carcinoma | 0,11 |
|  | Negative axilla | 0,11 |
|  | Partial breast irradiation | 0,11 |
|  | Breast-contour | 0,07 |
|  | Human breast carcinoma | 0,07 |
|  | Large breast cancers | 0,04 |
|  | Negative breast carcinoma cell lines | 0,04 |
|  | Non-obesity breast cancer | 0,04 |
|  | Obesity breast carcinoma | 0,04 |
|  | O-positive breast cancer | 0,04 |
|  | Positive breast cancer | 0,04 |
|  | Risk of breast cancer-related death | 0,04 |
|  | Multifocality | 0,04 |

Table S7. Category 7. Cancer. Original terms for early stage of Breast Cancer.

|  |  |  |
| --- | --- | --- |
|  | Term | % in the category |
|  | Axillary lymph node (ALN) | 2,23 |
|  | Axillary | 1,30 |
|  | Tumor size | 0,93 |
|  | Risk of recurrence | 0,87 |
|  | Tumor cells | 0,62 |
|  | ALNM Axillary lymph node metastasis | 0,49 |
|  | Tumor burden | 0,43 |
|  | Tumor-associated antigens (TAA) | 0,43 |
|  | Micrometastases | 0,37 |
|  | Solid tumors | 0,31 |
|  | Tumor expression | 0,31 |
|  | Basal(-)like | 0,25 |
|  | Gastric adenocarcinoma | 0,25 |
|  | Isolated tumor cells (itcs) | 0,25 |
|  | Tumor grade | 0,25 |
|  | Tumor-associated macrophages (tams) | 0,25 |
|  | Tumor-specific | 0,25 |
|  | Tumour-infiltrating cells (tiics) | 0,25 |
|  | Cancer progression | 0,19 |
|  | Macro-metastases | 0,19 |
|  | Macrometastasis | 0,19 |
|  | Tumor diameter | 0,19 |
|  | Tumor response | 0,19 |
|  | Tumor samples | 0,19 |
|  | Tumorigenesis | 0,19 |
|  | Invasive cancer | 0,12 |
|  | Intratumoral | 0,06 |
|  | Mammary carcinogenesis | 0,06 |
|  | Non-invasive cancer | 0,06 |
|  | Non-obesity cancer carcinoma | 0,06 |
|  | Novel oncogenic | 0,06 |
|  | Poor prognosis cancers | 0,06 |
|  | Positivity of tumor | 0,06 |
|  | Premalignant | 0,06 |
|  | Skeletal metastases | 0,06 |
|  | Sporadic cancers | 0,06 |