Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Ontological representation of the evolution of medical terminology related to the process of calcification in breast cancer

Aida Khakimova
Russian New University
Oleg Zolotarev
Russian New University
Shweta Agrawal
indore institute of science and technology
Ritu Tandon
SAGE University indore
Sanjiv Kumar Jain
Medicaps University
Swati Varun Yadav
Manipal Academy of Higher Education
Manish Varun Yadav
yadav.manish@manipal.edu

Manipal Academy of Higher Education

Article

Keywords: Ontology, medical terminology, breast cancer, calcification, temporal ontology. Temporal Ontology, Breast Cancer Calcification, Medical Ontology Engineering, Biomedical Terminology Evolution, Knowledge Representation in Healthcare, Ontology-Based Modeling

Posted Date: May 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9421423/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Ontological representation of the evolution of medical terminology related to the process of calcification in breast cancer

Aida Khakimova[1], Oleg Zolotarev[2], Shweta Agrawal[3], Ritu Tandon[4], Sanjiv Kumar Jain[5], Swati Varun Yadav[6*], Manish Varun Yadav[7*]

[1]Institute of Information Systems and Engineering Computer Technologies, Russian New University, Moscow, Russia, aida_khatif@mail.ru

[2]Institute of Information Systems and Engineering Computer Technologies, Russian New University, Moscow, Russia, ol-zolot@yandex.ru

[3]Indore Institute of Science & Technology, Indore, India, shweta.sagecse@gmail.com

[4]Institute of Advance Computing, SAGE University, bansalritu10@gmail.com

[5]Medicaps University, Indore, India, sanjivkumar.jain@medicaps.ac.in

[6&7*]Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India, yadav.swati@manipal.edu, yadav.manish@manipal.edu

Corresponding Author: Swati Varun Yadav[6] (yadav.swati@manipal.edu), Manish Varun Yadav[7] (yadav.manish@manipal.edu)

Abstract

A temporal ontology is created to describe the calcification process in breast cancer. A terminological comparison of three ontological descriptions of the calcification process is carried out over three time periods (2003-2009, 2010-2015, 2016-2021) considering the emergence of new terms. Candidate terms for inclusion in the ontology were programmatically selected from a collection of documents and approved by experts for inclusion in the ontology, considering their contextual meanings. Three ontological representations for describing the process of calcification contain concepts related to the process of calcification in breast cancer, have a hierarchy of 103 concepts, and are accompanied by a set of definitions that define these concepts. For each period, an ontology is built for term-like phrases for each period, and the relationships between terms are clarified based on the medical literature. The temporal ontology of terms for the calcification process revealed the most actively developing ontology branches.

Keywords: *Ontology, medical terminology, breast cancer, calcification, temporal ontology. Temporal Ontology, Breast Cancer Calcification, Medical Ontology Engineering,*

Biomedical Terminology Evolution, Knowledge Representation in Healthcare, Ontology-Based Modeling

I. Introduction

In recent years, various biomedical ontologies have emerged, such as the National Cancer Institute (NCI) Thesaurus [Golbeck et al. 2003], Foundational Model of Anatomy (FMA) [Rosse et al. 2003], Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT) [Schulz et al. 2011] etc. These ontology resources are widely used in various fields such as the standardization of medical data formats [Cimino et al. 2006], representation and integration of medical or clinical knowledge [Isern et al. 2012], and medical decision making [De Potter et al. 2012].

Biomedical terminology is marked by ambiguity and complexity due to the ongoing expansion of biomedical data; this further complicates intelligent biomedical applications. Furthermore, emerging biomedical ontologies are constructed autonomously, resulting in disparate approaches to designating identical biomedical components and, consequently heterogeneous issues [Wang et al. 2021].

Harmonization of terminology helps researchers and practitioners better understand terms. This coordination is essential in promoting interdisciplinary relationships and exchanging concepts and knowledge [Richesson et al. 2011]. Knowledge of terminologies is a very important aspect in the medical field, where the knowledge base may be very dynamic and changes continuously due to the occurrence of new diseases, viruses, genes, etc.

Objectives:

- Assess and document the ontological depiction of the calcification process in breast cancer while considering the terminology's evolution over time.
- Create an extensive ontology encompassing the conceptual organization and semantic connections between terms related to the calcification of breast cancer during three different periods (2003-2009, 2010-2015, 2016-2021).
- provide a common, easily accessible, and dynamic ontological resource for describing and cataloging terms that include the term calcification as they emerge as research into the calcification process in breast cancer.
- Provide opportunities for the exchange of medical information and scientific research.

The paper is organized into six sections: Introduction, Literature Review, Proposed Methodology, Results & Discussion, Comparative study, and Conclusion.

II. Literature Review

Sandun et al. proposed a mechanism to create a dynamic ontology for medical terms that keeps updating new relationships with the help of online and offline patient records [Sandun et al. 2017]. To extract domain-specific relationships, the approaches "Extracting gene-disease relationships", "Recognition of protein names using a reference book" [Tsuruoka et al. 2003], and "Recognition of biological objects based on a dictionary in the biomedical literature" [Yang et al. 2008] were used. There are ideas for biomedical object recognition systems that use machine learning and a web system [Sumathipala et al. 2015].

Many dynamic medical ontologies have been developed in past years. Some of them are Infectious Disease Ontology (IDO) [Infectious Disease Ontology 2020], Human Phenotype Ontology (HPO) [The Human Phenotype Ontology 2024], Translational Medicine Ontology (TMO) [Translational Medicine Ontology 2016], etc.

Drug knowledge for the presentation of entities & their relationship is presented [Yongjun et al. 2019]. The different knowledge bases represent the various activities: gene, drug, target, side effect, as well as connections: drug-gene, drug-drug, drug-disease, drug-target, drug-side effect [Tandon et al. 2020].

Large scale integration of knowledge base leads to many challenges: the main challenge involves mapping entities between two or more knowledge bases, another challenge relates to the integration of entity relationships. This challenge can be observed in existing drug knowledge bases, which demonstrate limited overlap [Ayvaz et al. 2015]. For example, some drug knowledge bases, such as PharmGKB [Whirl-Carrillo et al. 2012], primarily focus on drug-gene relationships, while others have other foci like SIDER drug side effects [Kuhn et al. 2016].

PubMed [NCBI Resource Coordinators 2018] is primarily utilized as the primary index for publications on biomedical research. for this purpose, curation requires a substantial amount of time and financial resources [Poux et al. 2014]; Curating a single protein can take up to a week and requires significant amount of human investment in both knowledge and effort [Poux et al. 2017].

Calcification is a pathological process marked by deposition of calcium. It is a regulated physiological process that occurs in bones and teeth. However, calcification is commonly found in soft tissues due to aging and various diseases [Proudfoot et al. 2019]. Calcifications may occur with benign changes. Yet significantly, they can be the initial signs of a malignant process in the mammary gland [Oksanchuk et al. 2016]. In some cases, calcifications stand out as the only symptom of breast cancer.

III. Proposed Methodology

A Python and PHP-based intelligent information system was created to conduct the analysis of scientific papers. The search was conducted by using the search terms (calcification [Abstract] AND ("neoplasms"[MeSH Terms] OR "neoplasms"[All Fields] OR "cancer"[All Fields])) AND ("breast"[MeSH Terms] OR "breast"[All Fields]) AND ("2003/01/01"[PubDate]: "2021/12/31"[PubDate]) in the PMC database, 833 articles were retrieved in which the term calcification appeared 2650 times.

The array of articles was divided into three collections - for 2003-2009, 2010-2015, 2016-2021. 57 articles (189 entries) were included from the collection for 2003-2009, the collection for 2010-2015 included 225 articles (700 entries) and the collection for 2016-2021 included 551 articles (1761 entries).

Based on the queries with specific keywords and the timeframe during which segments of scientific publications were retrieved from PubMed, information such as article titles, authors, abstracts, keywords, publication years, and article corpora were generated. Additionally, term-like expressions containing the term "calcification" were identified and emphasized. Terms and term-like phrases were highlighted programmatically in these collections.

A database is formed with the following fields: document title, author, abstract, year of publication, etc. Further, because of the automated processing of titles and abstracts of downloaded documents, terms and their contexts were identified, excluding stop words. Their expert verification was performed after the automatic selection of terms/term-like phrases.

The system automatically compares the selected term-like phrases with a list of traditional terms compiled by experts and based on electronic medical linguistic resources (MeSH, SNOMED CT, UMLS). This focuses on the possible new terms. To approve a potentially new term as a traditional one, an expert can view not only a list of potentially new terms, but also their contextual environments (neighborhoods). The extent by which one can exert control to more accurately define the term and its possibly new contextual meanings. In a similar way, an expert can work with a list of selected traditional terms to find their new meanings.

Phrases that were similar to the term "calcification" were found in each collection. Note that not every instance is a phrase that resembles a term. The number of term-like phrases is indicated further in the results section. For each time period, an ontology for term-like phrases was built using the expert method, and the relationships between entities were clarified on the basis of medical literature.

IV. Results & discussion

4.1 Ontology for the period from 2003 to 2009

Key terms and term-like phrases were identified (further in the text - "terms") from the collection of articles from 2003 to 2009. The set of terms obtained from this collection was divided by experts into 9 thematic clusters (Figure 1). To make topics clearer, the clusters were presented on different levels.

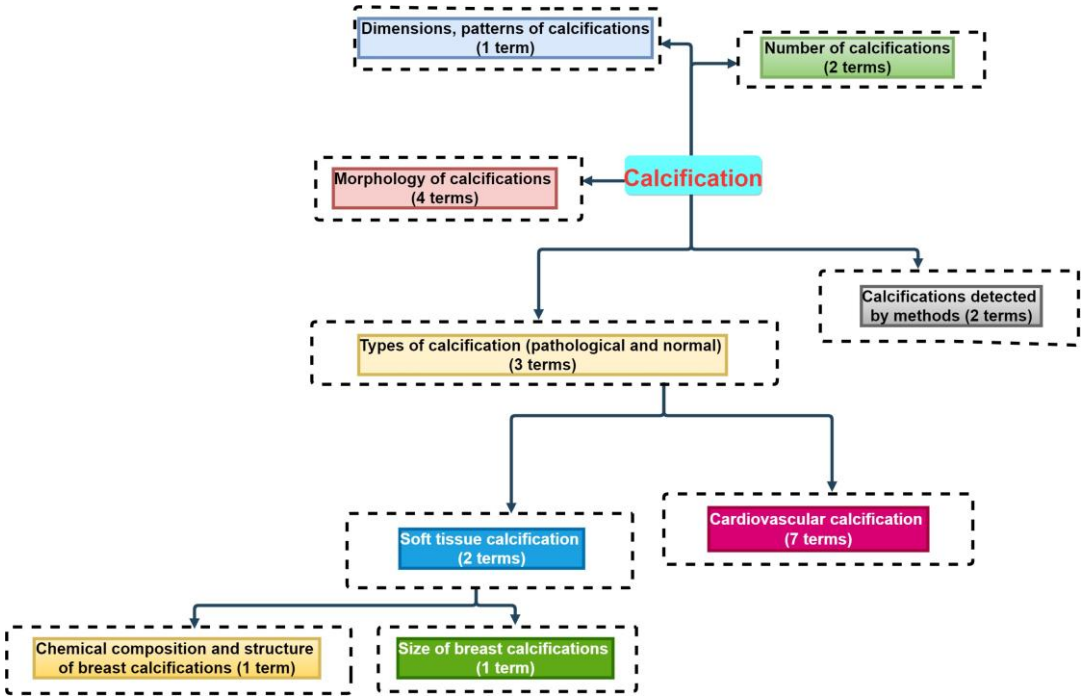


Figure 1. Clusters of terms identified from a collection of articles for the period from 2003 to 2009.

The first level identified clusters are: i) “Types of calcification (pathological and normal)” (3 terms); 2) “Morphology of calcifications” (4 terms); 3) “Calcifications detected by methods” (2 terms); 4) “Dimensions, patterns of calcifications” (1 term); 5) “Number of calcifications” (2 terms).

Clusters of the second level were identified for the “Types of calcification” cluster, based on the answer to the question “Where?” – These clusters are- “Soft tissue calcification” (2 terms) and “Cardiovascular calcification” (7 terms).

At the third level, additional classification was carried out in the "Soft tissue calcification" cluster, emphasizing the question "What are the characteristics of breast calcifications?" This

led to the emergence of two sub-clusters: "Size of breast calcifications" (one term) and "Chemical composition and structure of breast calcifications" (one term).

4.2 Ontology for the period from 2010 to 2015

Figure 2 shows the clusters from the collection of articles for the period 2010 to 2015. The terms were highlighted from this period and these new terms were added in the previously obtained 9 clusters for enrichment. New clusters of all levels were then identified.

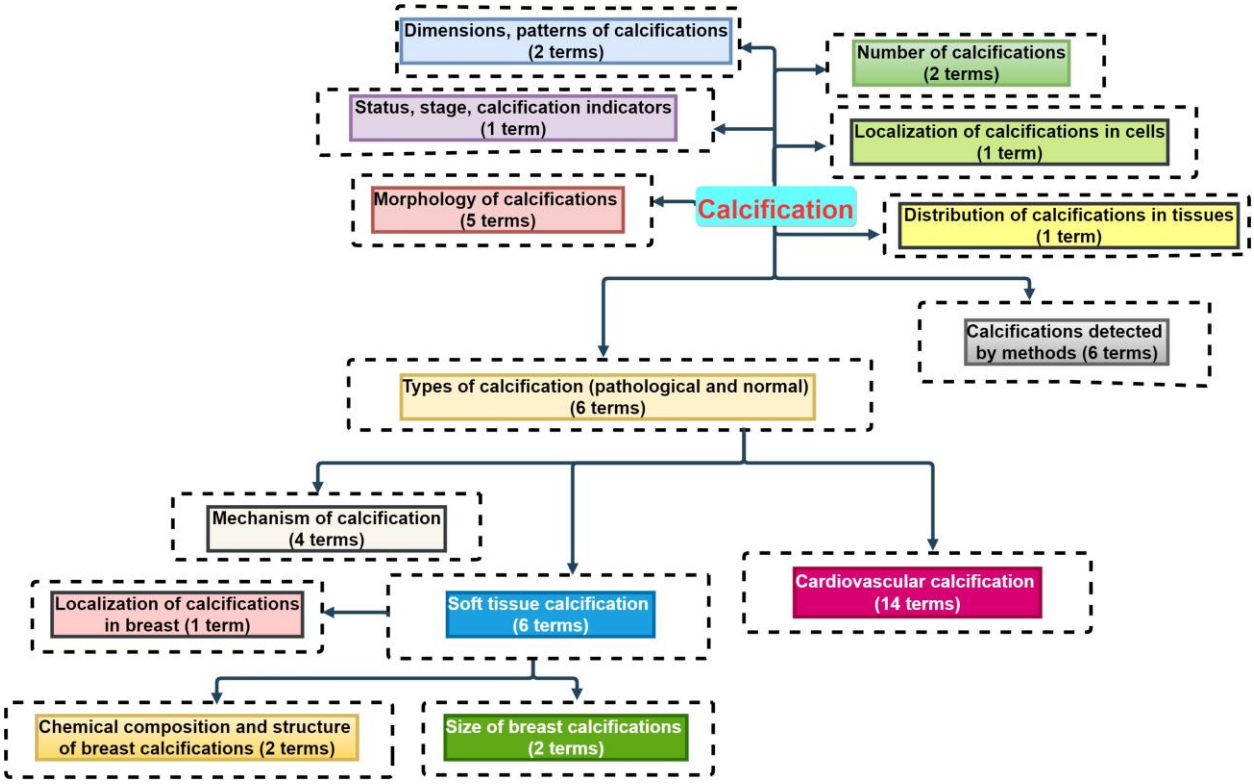


Figure 2. Clusters of terms identified from a collection of articles for the period from 2010 to 2015.

New terms appeared in the previously identified clusters of the first level are: 1) “Types of calcification (pathological and normal)” (3 terms) ; 2) “Morphology of calcifications” (1 term) ; 3) “Calcifications detected by methods” (4 terms); 4) “Dimensions, patterns of calcifications” (1 term). In the “Number of calcifications” cluster, the number of terms were not changed.

New first-level clusters were discovered by experts, including "Status, stage, calcification indicators" (1 term), "Localization of calcifications in cells" (1 term), and "Distribution of calcifications in tissues" (1 term).

Two new phrases, "Soft tissue calcification" (5 terms) and "Cardiovascular calcification" (7 terms), appeared in the previously established second-level clusters. The response to the query "How?" led to the identification of a new second-level cluster for the "Types of calcification" cluster, which is known as the "Mechanism of Calcification" cluster (4 words).

In the previously identified clusters of the third level, new terms appeared: “Chemical composition and structure of breast calcifications” (1 term) and “Size of breast calcifications” (1 term). For the “Soft tissue calcification” cluster, a new third level cluster was identified based on the answer to the question “Where is the location in the chest?” – This led to the identification of the cluster “Localization of calcifications in the mammary glands” (1 term).

4.3 Ontology for the period from 2016 to 2022

Figure 3 shows the clusters from the published articles from 2016 to 2022, the terms were highlighted. During this period, the previously obtained 14 clusters were enriched with new terms.

In the previously identified clusters of the first level, new terms appeared: 1) “Types of calcification (pathological and normal)” (1 term); 2) “Morphology of calcifications” (5 terms); 3) “Calcifications detected by methods” (5 terms); 4) “Dimensions, patterns of calcifications” (3 terms); 5) “Distribution of calcifications in tissues” (1 term); 6) “Localization of calcifications in cells” (3 terms); 7) “Status, stage, calcification indicators” (2 terms). In the “Number of calcifications” cluster, the number of terms were not changed.

Experts identified a new first-level cluster “Induced calcification.” The “Uncertain Cluster” includes 6 terms that are difficult to attribute to the identified clusters. New terms appeared in the previously identified second-level clusters are: “Soft tissue calcification” (2 terms); “Cardiovascular calcification” (2 terms); “Mechanism of calcification” (1 term). In the previously identified clusters of the third level, new terms appeared: “Chemical composition and structure of breast calcifications” (5 terms): “Size of breast calcifications” (1 term); “Localization of calcifications in the mammary glands” (1 term).

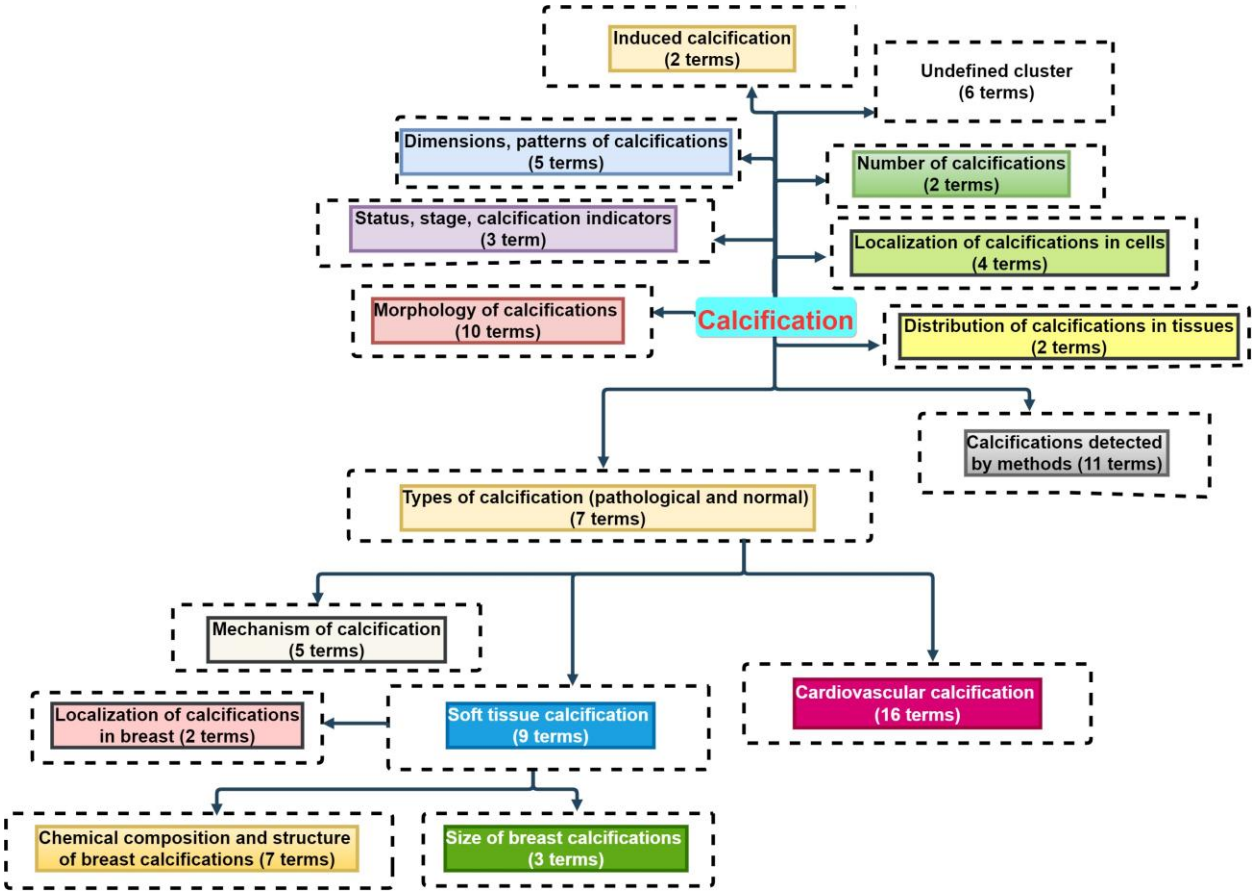


Figure 3. Clusters of terms identified from a collection of articles for the period from 2016 to 2022.

The complete ontology is shown in Figure 4. For 2003-2009, total 30 terms were included in the ontology, for 2010-2015 - 58 terms, for 2016-2021 - 94 terms were included. The total 94 terms were included. Based on medical literature, experts generated significance for the terms on the list. Database updates regarding definitions have been created.

4.4 Discussion

The most actively developed clusters are “Sizes, patterns of calcifications”, “Morphology of calcifications”, “Chemical composition and structure of breast calcifications”.

4.4.1 Cluster “Morphology of calcifications”

In ontology for 2003-2009, let’s consider the cluster “Morphology of calcifications” (Figure 5) that Includes 4 terms - calcification morphology, amorphous calcifications, Benign Calcifications, suspicious calcifications. In 2010-2015 the term popcorn-like calcification was added. In 2016-2022 this cluster included 5 more terms - pleomorphic calcifications, Coarse Heterogeneous microcalcifications, eggshell calcification, punctate calcification,

cancer-related calcifications (10 in total). From 2010 to 2015, the phrases "amorphous calcifications" and "beneign calcifications" were not used; they reappeared from 2016 to 2022.

.

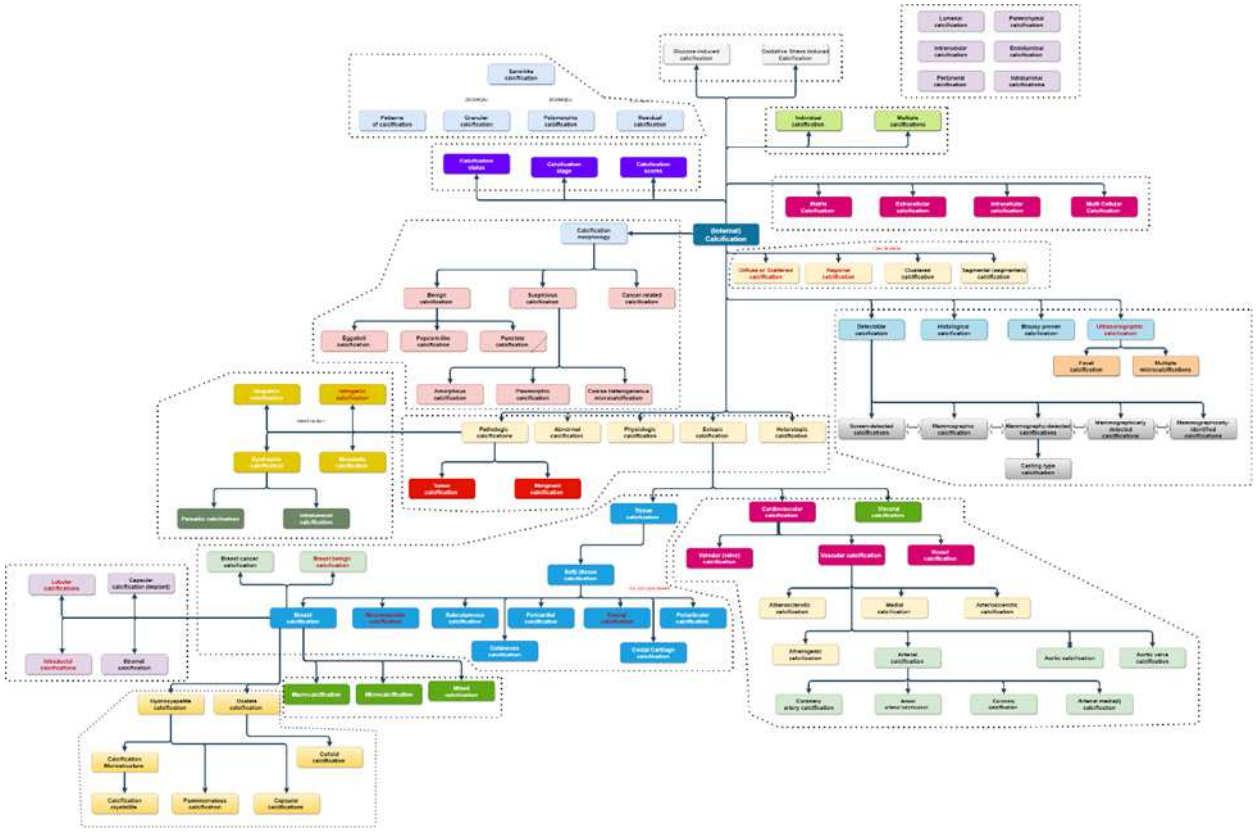


Figure 4. Breast cancer calcification ontology from 2003 to 2021.

The form of calcifications serves as the most significant factor in the distinction between benign and malignant calcifications, and that is the reason why calcification morphology is a generic term (1st order branch). Benign calcifications and suspicious calcifications, respectively, refer to second-order branch calcifications that are benign and cause for concern. The term "cancer-related calcifications" was introduced to the second level in 2016–2022. This was followed by third-order terms denoting varieties of both benign and suspicious classifications. Varieties of benign calcifications are eggshell calcification, popcorn-like calcification, and punctate calcification [Breastcancer.org 2024]. Amorphous calcifications were classified as types of suspicious calcifications [Radiopaedia 2024]. Most amorphous calcifications are assessed as BI-RADS 4B - suspicious, requiring biopsy [Dev et al. 2022]. Pleomorphic calcifications can sometimes signal the presence of ductal carcinoma in situ (DCIS), with or without an invasive breast cancer present as well. Amorphous, coarse

heterogeneous, fine linear or fine-linear branching, and fine pleomorphic calcifications are the four types of calcifications with suspicious morphology that fall under the category of pleomorphic calcifications. Together with amorphous microcalcifications, coarse heterogeneous microcalcifications are considered as being of intermediate statistical significance [Radiology Assistant 2024].

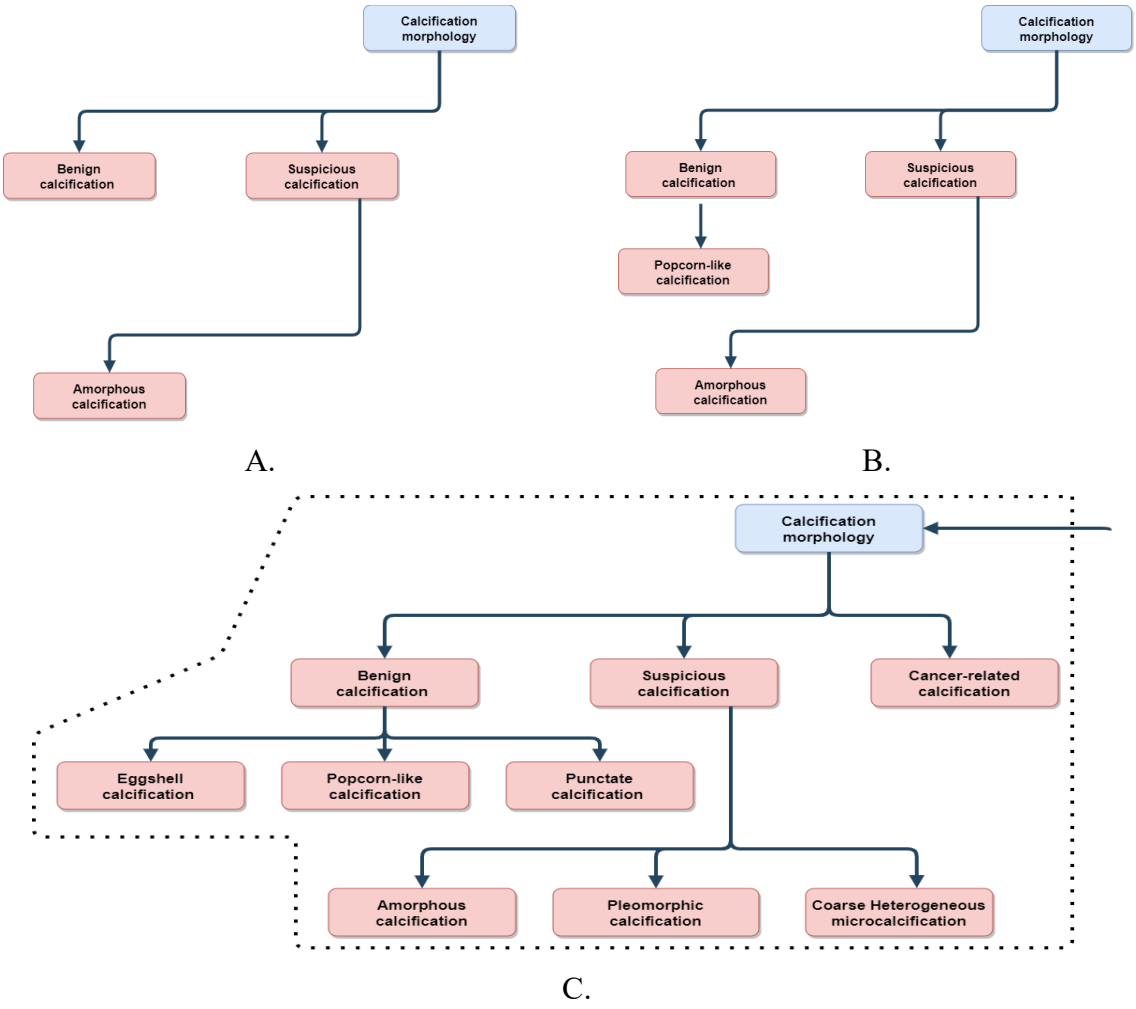


Figure 5 (A, B, C). Evolution of the Calcification Morphology cluster of the Breast Cancer Calcification Ontology from 2003 to 2021.

4.4.2 Cluster "Chemical composition and structure of breast calcifications"

In the cluster "Chemical composition and structure of breast calcifications" in 2003-2009 one term Psammomatous calcifications was assigned (was not found in 2010-2015 and reappeared in 2016-2021). In 2010-2015 the terms Capsular calcification and hydroxyapatite calcification appeared. Cluster for 2016-2022 shown in Figure 6.

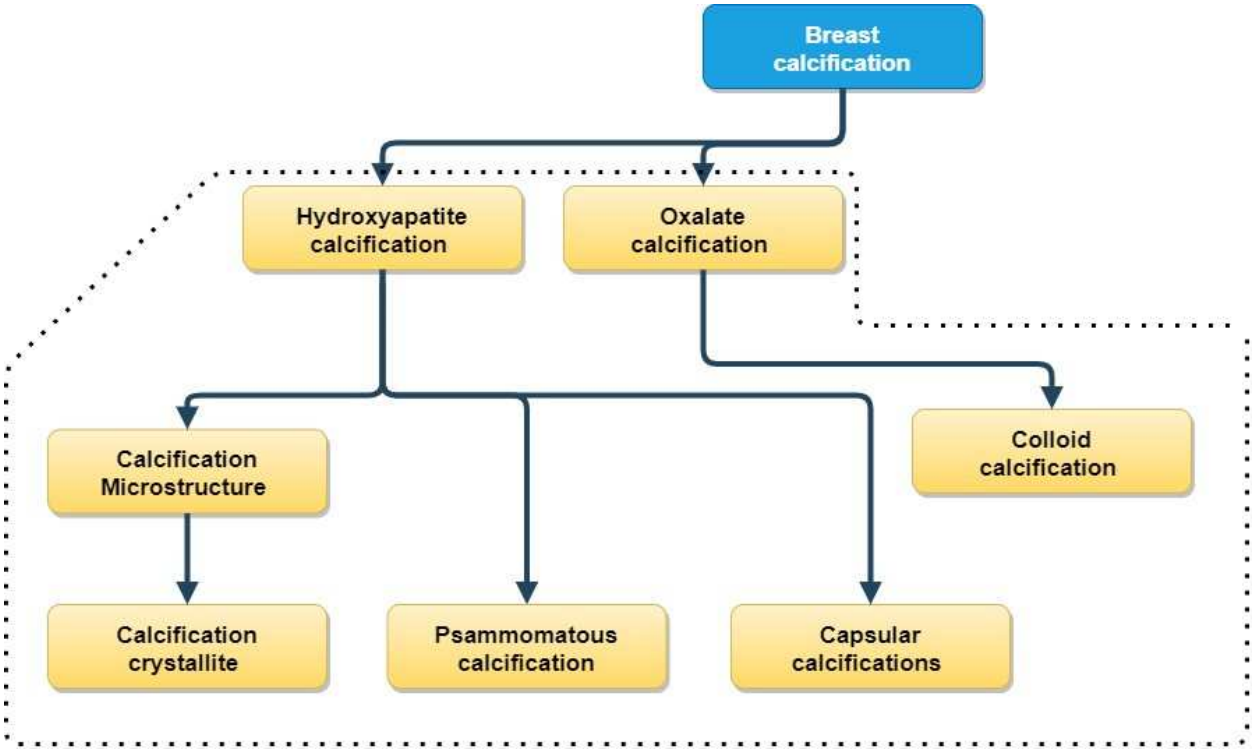


Figure 6. Cluster "Chemical composition and structure of breast calcifications" for the period from 2016 to 2021.

Hydroxyapatite calcification emerged between 2010 and 2015, while oxalate calcification gained recognition from 2016 to 2021. Other terms were categorized based on their composition. Colloid calcification consists of calcium oxalate [Guerlain et al. 2019]. The calcification microstructure, specifically hydroxyapatite microcalcification, encompasses calcification crystallites [Gosling et al. 2019]. Capsular calcification is characterized by the presence of calcium phosphate [Guerlain et al. 2019]. Psammomatous calcifications refer to deposits of hydroxyapatite crystals with a layered psammomatous appearance [Suresh et al. 2013].

4.4.3 Cluster "Cardiovascular calcification"

The Cardiovascular calcification cluster experienced significant development between 2010 and 2015. During the initial time, experts categorized 7 terms into this cluster. This number increased to 14 in the subsequent period, and in the third period, it further increased to 16 (Figure 7).

The terms of the following level are as follows: vascular calcification is also known as vessel calcification, and valvular calcification is also referred to as valve calcification. These terms specify the specific location of calcifications in major blood vessels (Arterial calcification, aortic calcification, and Aortic Valve Calcification).

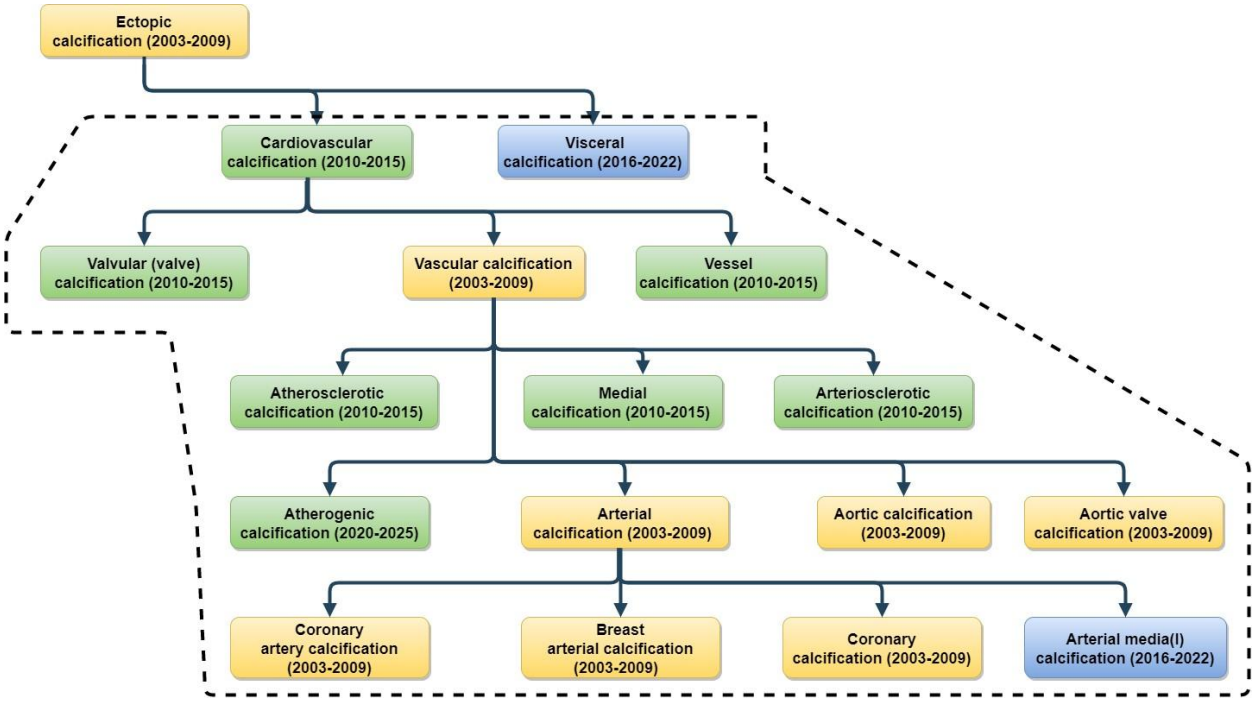


Figure 7. Cardiovascular calcification cluster. (Yellow color – terms assigned to the cluster for the period 2003-2009, green – 2010-2015, blue – 2016-2022).

Arterial calcification is classified into two categories based on the location of the calcifications: intimal calcification, which occurs in the inner membrane, and medial calcification, which occurs in the middle membrane. Another division of the ontology pertains to the localization of calcification in the intima (Atherosclerotic Calcification) and media (Arteriosclerotic Calcification, medial calcification). The words breast arterial calcification, coronary artery calcification, coronary calcification, and arterial media(l) calcification are all classified forms of arterial calcification by experts [Suh et al. 2018].

The phrase "atherogenic calcification" refers to a distinct branch. Atherogenic variables are factors that, when they affect the body, lead to the development of atherosclerosis. There are specific phases of calcium deposition that may have a greater probability of encouraging the development of atherosclerosis [Shi et al. 2020]. Also, a separate branch is formed by the term Visceral Calcification. Visceral calcification can be observed in various organs, including heart valves, myocardium, kidneys, lungs, and brain.

4.4.4 Cluster “Calcifications detected by methods”

The cluster of calcifications is recognized using actively developed procedures across all time periods. In the beginning, experts categorized 2 terms into the cluster. In the second period, the number increased to 6, and in the third period, it further increased to 11. (Figure 8).

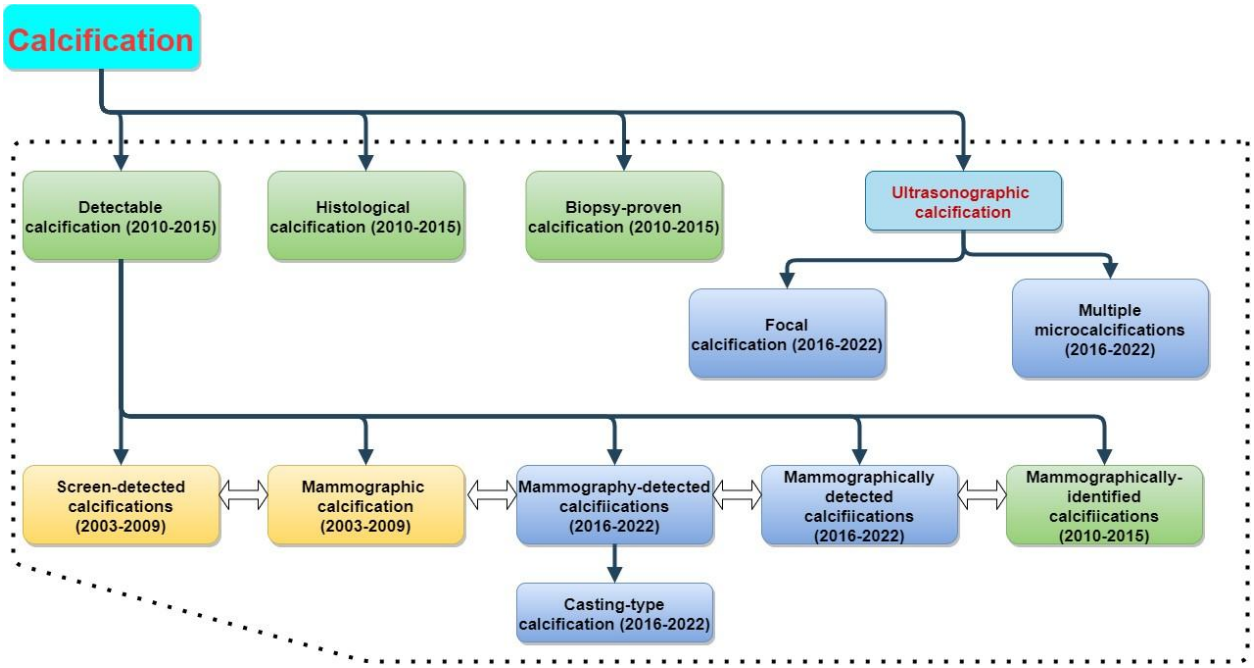


Figure 8. Cluster Calcifications detected by methods. (Yellow color – terms assigned to the cluster for the period 2003-2009, green – 2010-2015, blue – 2016-2022).

The initial set of words pertains to calcifications identified using various techniques, including detectable (micro) calcifications, microcalcifications verified by biopsy, and histological calcification. All terms were present during the years 2010 to 2015.

The second category refers to the identification of calcifications using mammography, sometimes known as mammographic calcifications. Mammographic microcalcifications are regarded as initial indicators of breast cancer (BC). Various synonyms are employed in this context: mammographically-identified calcifications (2010-2015), mammographically detected calcifications (2016-2022), and Mammography - detected calcifications (2016-2022). The term "screen-detected calcifications" was utilized throughout the period of 2003-2009. The presence of casting-type calcification is strongly associated with axillary node metastases, absence of estrogen receptor expression, and overexpression of HER2. Therefore, casting-type calcification is considered an autonomous prognostic determinant for patients with breast cancer [Li et al. 2019].

The words "focal calcifications" and "multiple microcalcifications" emerged between 2016 and 2022. The identification of these types was accomplished by making use of ultrasonography. It was hypothesized that the word "ultrasonographic calcification" should be positioned at the highest level within the cluster.

4.4.5 Cluster “Types of calcification (pathological and normal)”

The cluster of Calcification Types (pathological and normal) experienced substantial expansion from 2010 and 2015. In the beginning, experts categorized 3 terms into the cluster. Futher the number increased to 6, and in the third period, it further increased to 7 (Figure 9).

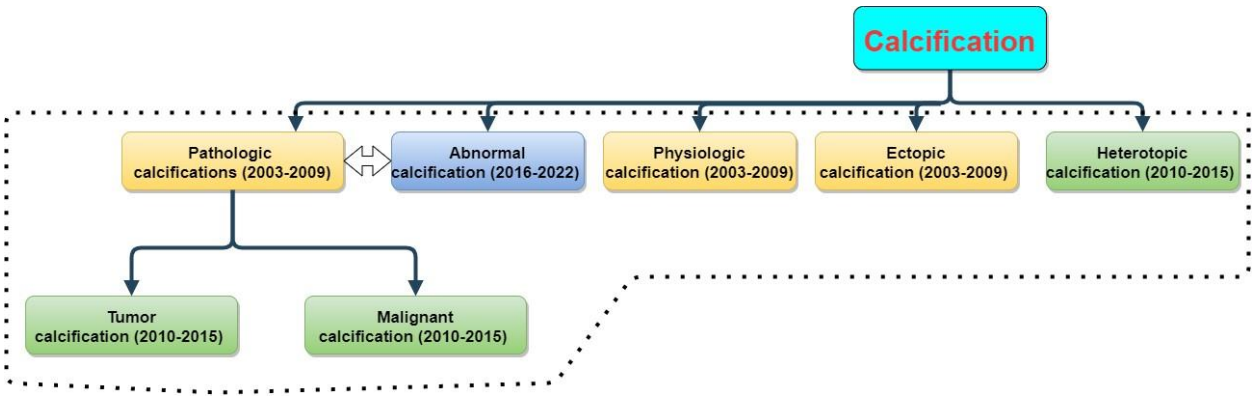


Figure 9. Cluster Types of calcification (pathological and normal). (Yellow color terms assigned to the cluster for the period 2003-2009, green color terms assigned to the cluster for the period 2010-2015, and blue color terms assigned to the cluster for the period 2016-2022).

Between 2003 and 2009, the phrases pathological calcification (refers to improper deposition of calcium in soft tissues), physiologic calcifications, and ectopic calcifications (describe incorrect biomineralization in soft tissues) were introduced [Giachelli et al. 1999]. The term Heterotopic calcification was included in this level of definition from 2010 and 2015. There were introductions of criteria for pathological calcification, including malignant calcifications and tumor calcification. During the period from 2016 to 2022, the phrase abnormal calcification was employed interchangeably with pathological calcification.

4.4.6 Cluster “Calcification of soft tissues”

The Soft tissue calcification cluster experienced significant development between 2010 to 2015. In the beginning, experts categorized 2 terms into the cluster. In the second period, the number increased to 6, and in the third period, it further increased to 9 (Figure 10).

The phrase "tissue calcification" originated between 2010 to 2015, along with the more specific term "soft-tissue calcification." The localization of calcification allowed for the distinction of the following categories: breast microcalcifications and Breast Cancer microcalcification (2003-2009); Cutaneous Calcification and Subcutaneous calcification (2010-2015); pericardial calcification and periarticular calcification (2016-2022); Costal

Cartilage Calcification (2016-2022). The terms breast benign calcification, neurovascular calcification, and fascial calcification were anticipated.

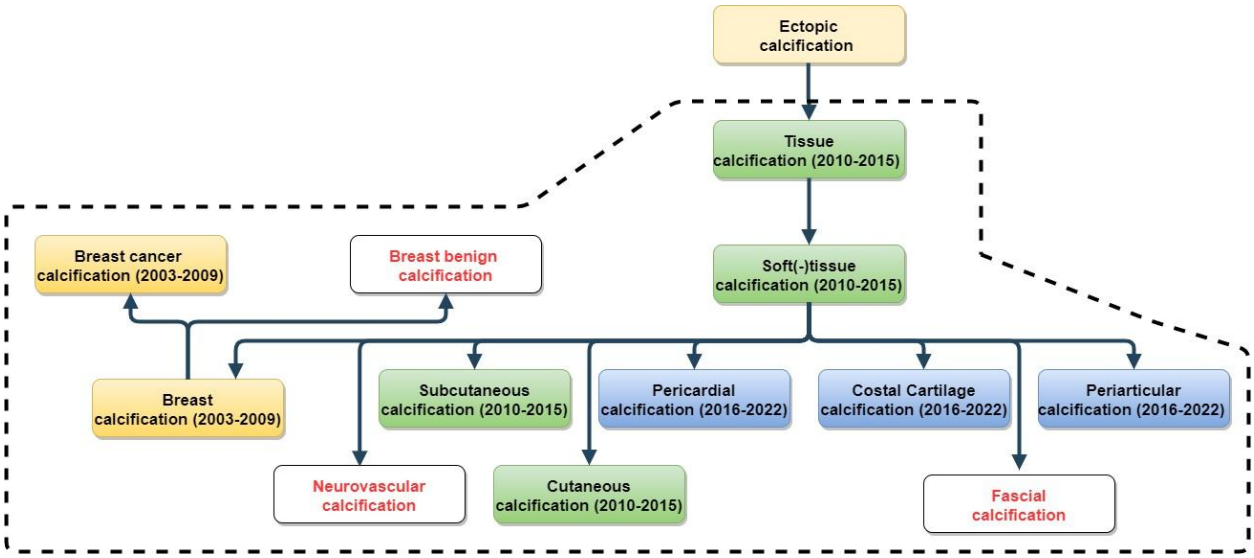


Figure 10. Soft tissue calcification cluster. (Yellow color – terms assigned to the cluster for the period 2003-2009, green – 2010-2015, blue – 2016-2022).

4.4.7 Small clusters identified in the period from 2003 to 2009.

A cluster of calcifications developed between the years 2003 to 2009. The presence of individual and multiple calcifications has remained unchanged. The breast calcification size cluster emerged between 2003 to 2009. The two terms are microcalcification and macrocalcification. The term "mixed calcification" was introduced in the years 2016-2022 to describe a pattern of calcifications that includes both microcalcifications and macrocalcifications [Bianchi et al. 2019].

The cluster "Dimensions, patterns of calcifications" in 2003-2009 encompassed a single phrase, specifically "Patterns of calcification." The term "Sand-like calcification" was introduced between 2010 and 2015. From 2016 to 2021, this cluster encompassed 5 terms: granular calcification, residual calcifications, and polymorphic calcification. During the period from 2016 to 2022, the term "sand-like calcification" was not used.

Small clusters identified in the period from 2010 to 2015.

The mechanism of calcification cluster appeared in 2010-2015 (4 terms). In 2016-2022 a new cluster term appeared (Figure 11). There are two processes by which pathologic calcification occurs: metastatic and dystrophic [Yu et al. 2017]. Dystrophic calcification refers to the

accumulation of calcium salts in deteriorated tissues, without any disruption in the overall mineral balance of the body.

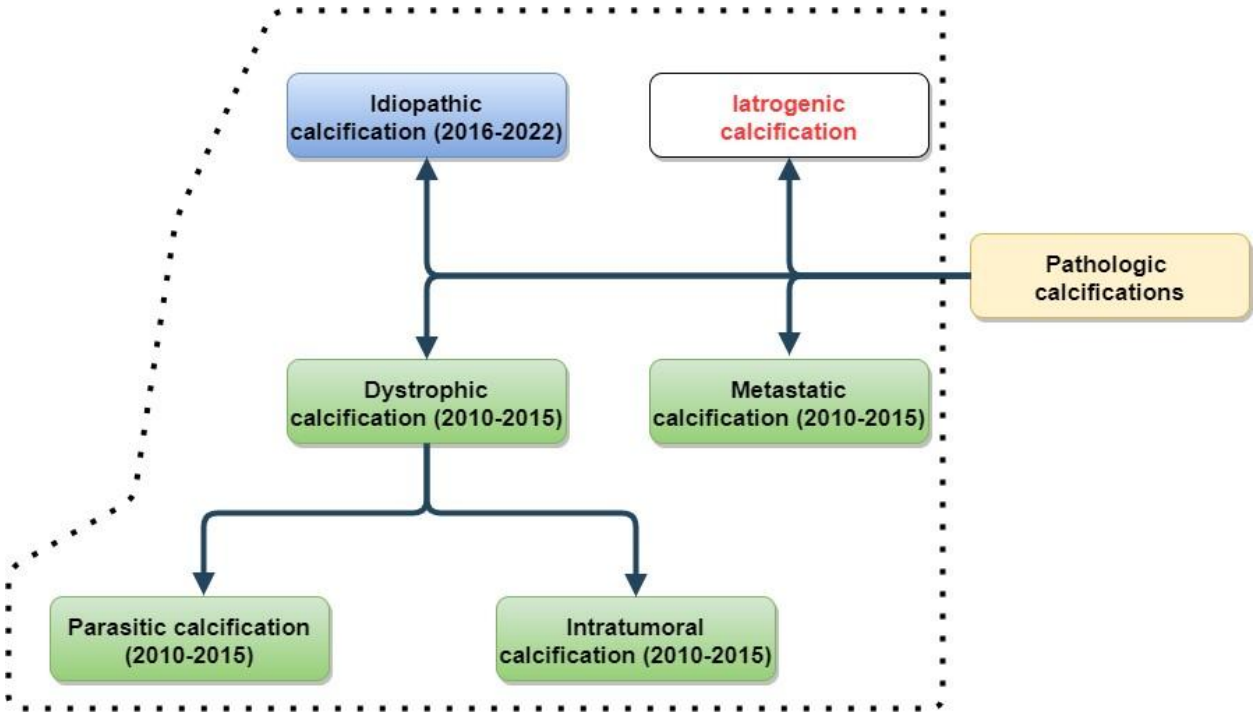


Figure 11. Calcification mechanism cluster. (Green color – terms assigned to the cluster for the period 2010-2015, blue – 2016-2022. Red font is a potential term).

Metastatic calcification refers to the accumulation of calcium salts in tissue that is otherwise healthy. Metastatic calcification with hypercalcemia refers to the deposition of calcium in tissues that were previously healthy, while dystrophic calcification occurs in tissues that were previously injured [Senba et al. 2011].

The term "soft-tissue calcifications" is classified into four categories based on the process of formation: dystrophic, iatrogenic, metastatic, and idiopathic. These categories are determined by considering the clinical and biochemical association [Faria et al. 2019]. The phrase idiopathic calcification was introduced during the years 2016 to 2022. The phrase iatrogenic calcification was included as a possible contender.

Intratumoral calcification is a type of pathologic calcification usually considered to be dystrophic calcification caused by degenerative changes of the tissue, such as necrosis or hemorrhage, except for a few bone-forming tumors [Dellamonica et al. 1977]. Parasitic calcifications are heterotopic dystrophic calcifications, their late apparition in the evolution of

the parasitic disease is not subordinate to any metabolic dysfunction [Radiology assistant 2023].

The phenomenon of calcifications clustering in the mammary glands was observed between 2010 to 2015. The phrase "capsular calcification" refers to the calcification that occurs around an implant. The phrase "breast calcification" refers to the presence of calcium deposits in the breast tissue. A new cluster term, stromal calcification, based on breast tissue, was introduced between 2016 to 2022. The potential possibilities for the cluster include lobular calcifications and intraductal calcifications, which are determined based on their location in the breast tissue.

The cluster's status, stage, and calcification signs emerged between 2010 to 2015. The phrase "Calcification score" refers to a numerical value that quantifies the amount of calcification present. During the period of 2016-2022, two additional terms were introduced, namely "calcification status" and "calcification stage". The phenomenon of Calcifications in Cells cluster was observed during the years 2010 to 2015 (1 term: Matrix calcification). Between 2016 to 2022, three additional terminologies were introduced: extracellular calcification, intracellular calcification, and multi-cellular calcification.

The occurrence of calcifications in tissues exhibited a clustering pattern between the years 2010 to 2015 (Calcification that is grouped together). Between 2016 to 2022, an additional term, segmental (segmented) calcification, was included.

The BI-RADS atlas provides the following descriptions for the distribution of calcifications (1): Clustered: There are at least 5 calcifications that are concentrated in a small area of tissue. (Linear: The calcifications are arranged in a straight line, indicating that they may be deposits in a duct.) Segmental: Calcium deposits found specifically in the ducts and branches of a segment or lobe [Egshatyan et al. 2017]. The words Diffuse or dispersed calcification and regional calcification, as defined in the BI-RADS atlas, were suggested as potential options for the cluster.

4.4.8 Small clusters identified in the period from 2016 to 2022.

The occurrence of the Induced calcification cluster was seen during the years 2016 to 2022. The study encompassed two distinct words, namely glucose-induced calcification and oxidative stress-induced calcification.

The cluster had the following 6 terms: luminal calcification, parenchymal calcification, intranodular calcification, endoluminal calcification, peripheral calcification, and intraluminal calcification due to the absence of relevance soft tissue or the circulatory system.

V. Comparative Study

5.1 Comparison of the ontology with existing ontologies.

The term calcification has been studied in existing ontologies. The simplified Russian-language nomenclature of calcification, derived from Lanzer (Figure 12). The taxonomy consisted of 17 terms organized in a hierarchical structure: categories - biological calcification, physiological calcification, ectopic calcification; tissues - bones, soft tissues/cartilage, cardiovascular; localization - valves, vessels, intimal, medial; risk factors and so on.

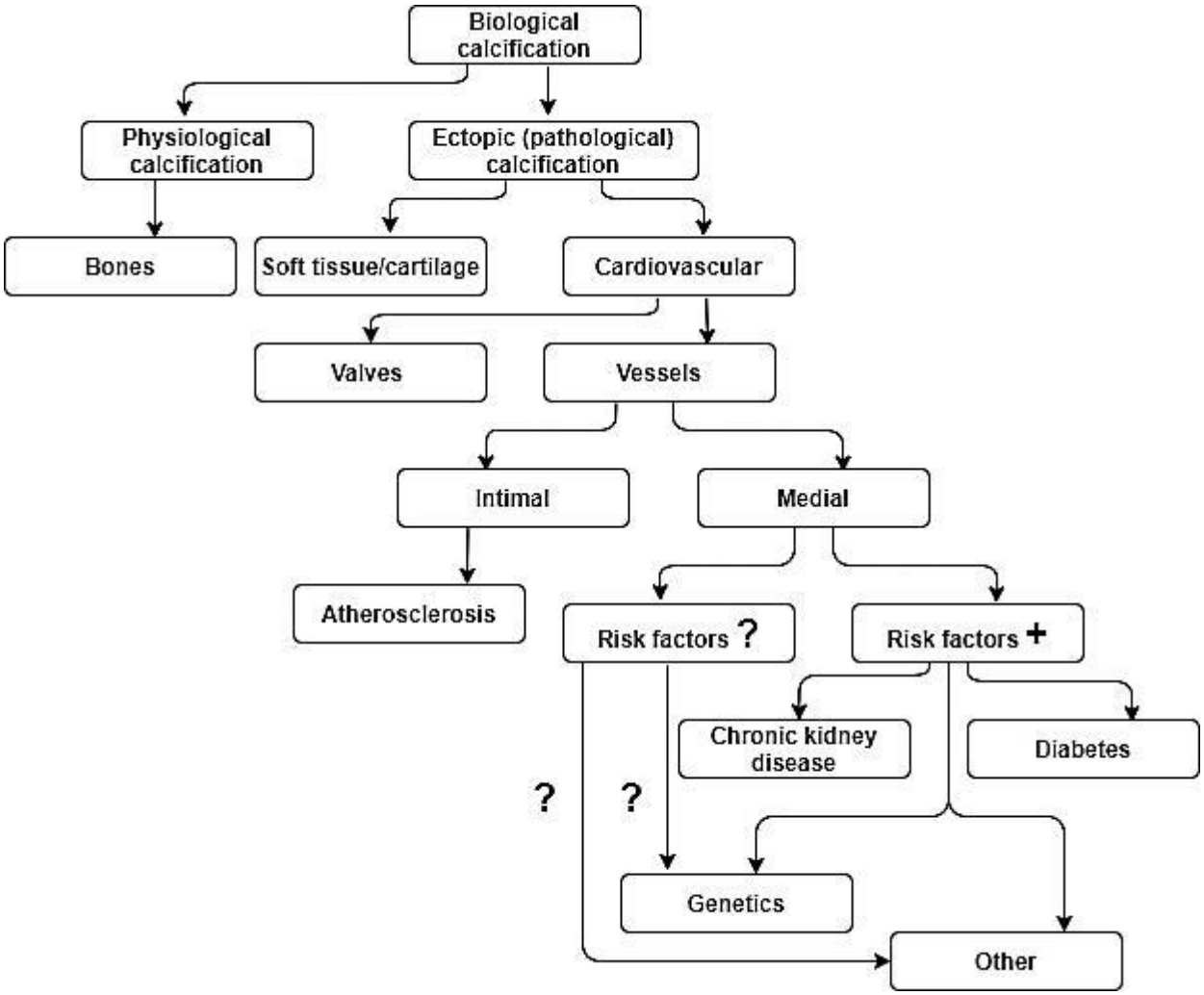


Figure 12. Simplified nomenclature for calcification (translated into English). Adapted from R. Lanzer [Lanzer et al. 2014].

The Rat Genome Database [Ontology browser 2024] provides a comprehensive analysis of the Ontology for calcification, which defines calcification as a disorder of calcium metabolism caused by the accumulation of calcium deposits in soft tissues. The ontology comprises a total of 7 terms. Illness Calcinosis is a type of acquired metabolic disease that

falls under the category of nutritional and metabolic diseases. It specifically affects calcium metabolism and can be classified as a mineral metabolic disease.

This resource uses data from 19 ontologies. Data are accessed through NCBI PubMed. Data can be retrieved from the database using the link search page or links on the object page. An ontology for pathological calcification in SNOMED CT is reviewed [3]. The ontology includes 11 terms: body structure, morphologically altered structure, morphologically abnormal structure, lesion, degenerative abnormality, heterotopia, deposition, lesion of degenerative abnormality, pathologic mineralization, acquired heterotopic structure, pathologic calcification. These ontologies are not appropriate for our terms in terms of their content.

SNOMED CT provides extensive data analytics. The disadvantages of SNOMED CT include the rather complex organization for accessing data, but it allows to extract and add a wide variety of data from/to the system. An example of visualization of the hierarchical environment of a term is shown in Figure 13:

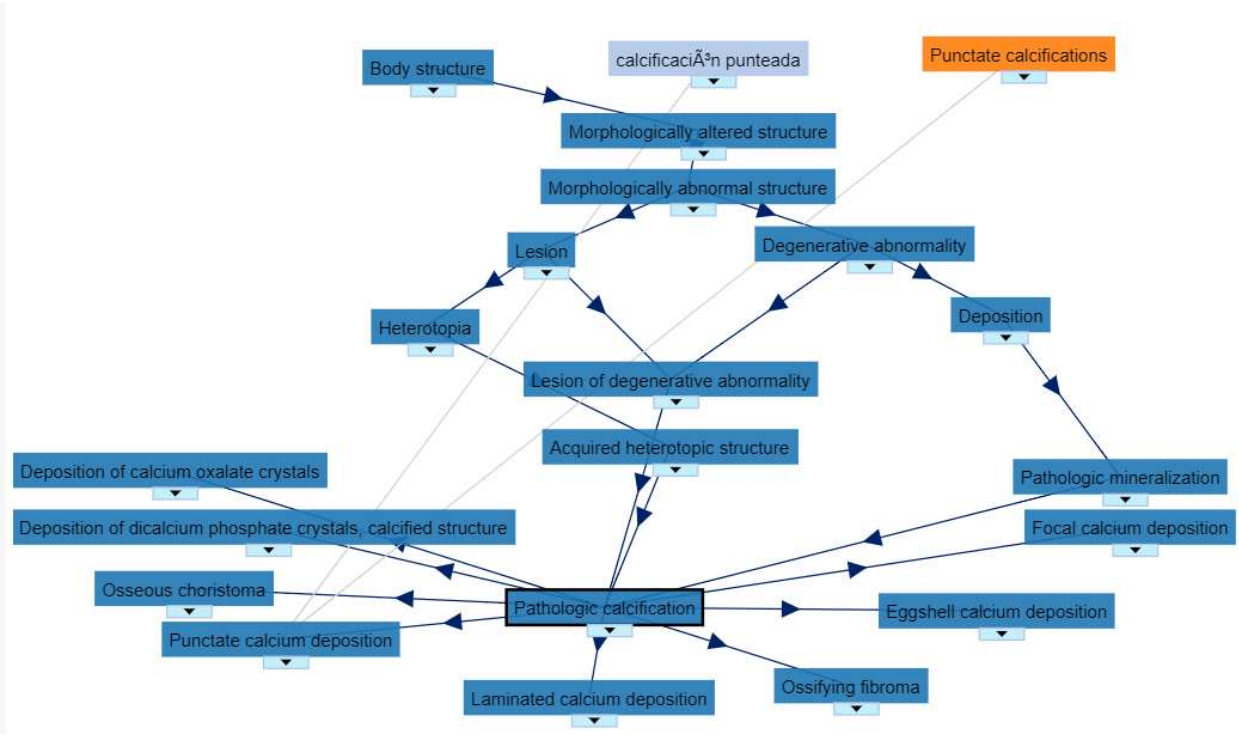


Figure 13. Visualization of a “pathologic calcification” term's partial environment.

RGD and SNOMED CT resources have a complex structure and use different sources, ontologies, and visualization methods. Both resources provide very comprehensive methods for accessing data.

The resources described above represent extensive structures and methods for accessing data, but this work does not address the issue of building a competing ontology, but rather the development of a mechanism using modern ontologies to search for new terms in medicine, which is not typical for any of the resources described above.

Our system is a superstructure over other systems and solves specific problems of new terminological knowledge construction based on known resources. In addition to highlighting new medical terms, the system allows to build trends in new terminology changes depending on time, which makes it possible to determine both increasing and decreasing trends.

5.2 Comparison of ontology terms with medical resources

The search used the titles and abstracts of papers. Following an automated search for instances of the term "calcification" (yielding a total of 2,658 occurrences, although not all of them constituted term-like phrases) experts conducted a semantic analysis of the surrounding contexts to identify term-like sentences that incorporated the term "calcification". Experts discovered 94 term-like phrases during a semantic study of 2658 occurrences spanning from 2003 to 2022.

The MeSH thesaurus [National library of medicine 2023] was used to check the dictionary status of 94 term-like phrases found by experts in the articles. In MeSH, 4 terms with definitions were found: Calcification, Physiologic (Year introduced: 1965), Vascular Calcification (2012). Among the Supplementary Concept the following terms were found: Arterial calcification of infancy (2010), Aortic Valve, Calcification of (2012). These terms are included in our ontology. As a result of the check, only 4 out of 94 term-like phrases coincided with MeSH thesaurus terms that are they have dictionary status.

The SNOMED CT thesaurus was used in the following step to determine if 94 term-like phrases from the ontology were included in dictionaries [Bioportal SNOMED CT 2023]. In SNOMED CT, 16 words were identified: Arterial calcification, Punctate calcification, valvular calcification, vascular calcification, individual calcification, metastatic calcification, myocardial calcification, periarticular calcification, pericardial calcification, dystrophic calcification, ectopic calcification, idiopathic calcification, and cardiovascular calcification. Following the check, only 15 of the 94 term-like sentences in the ontology had dictionary status—that is, they coincided with terms found in the SNOMED CT thesaurus.

52 terms were located in UMLS: "Popcorn" calcification; calcification of the abdominal aorta; amorphous calcification; calcification of the aortic arch and valve; calcification of the arteries; calcification of the arteries; calcification of the breast; benign and widespread breast

calcification; cluster of calcification; The conditions that can cause calcification of cartilage, coronary artery calcification, pulmonary valve calcification, cardiac valve calcification, heterogeneous calcification, idiopathic basal ganglia calcification, individual calcification, intra-abdominal calcification, intracranial arterial calcification, diffuse intracranial calcifications, dystrophic calcification, ectopic calcification, eggshell radiographic calcification, and extracranial calcifications include focal lung calcification, heart valve calcification, pulmonary valve calcification, and cardiac valve calcification. Intraluminal calcification on plain abdominal radiographs or prenatal ultrasound; Mammographic calcification and others. Only 28 of the 94 term-like phrases in the ontology that were found to have dictionary status because of the check coincided with terms found in the UMLS thesaurus.

VI. Conclusion

The proposed research presented the importance of ontological modeling and temporal analysis in understanding the calcification process in breast cancer. A robust ontological framework through careful selection of candidate terms and expert validation has been created. The framework not only captures the breadth of knowledge surrounding breast cancer calcification but also accounts for its evolution over time. By conducting a temporal analysis across three distinct time periods, the dynamic nature of medical terminology is also being analyzed. The proposed framework provides valuable insights into the emergent trends and developments within the field and will serve as a valuable resource for researchers, clinicians, and policymakers alike, facilitating deeper understanding, more precise communication, and, ultimately, more effective interventions in the fight against breast cancer.

Declarations:

Ethical Approval: Not applicable

Consent to Participate: Not applicable

Consent to Publish: On behalf of all authors, the corresponding author states that all are agreed for publication of this article

Authors Contributions:

Aida Khakimova: Paper drafting, methodology, discussion.

Oleg Zolotarev: Proof reading, figure drafting and literature review.

Shweta Agrawal: Proof reading, methodology, discussions.

Sanjiv K. Jain: Literature review and discussion.

Ritu Tandon: Paper drafting and Literature review.

Swati Varun Yadav: Proof reading, methodology.

Manish Varun Yadav: Data Collection, writing and drafting.

Funding : No funding was received for this research.

Competing Interests: Authors do not have any financial or non-financial interests that are directly or indirectly related to the work submitted for publication.

Availability of data and materials: Not applicable

References

[1]. Ayvaz S, Horn J, Hassanzadeh O, et al. Toward a complete dataset of drug-drug interaction information from publicly available sources. J Biomed Inform 2015;55:2017.

[2]. Bioportal SNOMED CT (2023), Systematized nomenclature of medicine clinical terms https://bioportal.bioontology.org/ontologies/SNOMEDCT. Accessed on 20 Dec 2023

[3]. Breastcancer.org.(2024) Accessed on 27 Jan 2024

[4]. Cimino JJ, Zhu X. The practical impact of ontologies on biomedical informatics. Yearbook of medical informatics. 2006;15(01):124-35.

[5]. D. Bianchi, U. Morandi, A. Stefani, B. Aramini. A case of mixed-pattern calcifications in multinodular goiter associated with a benign diagnosis, International Journal of Surgery Case Reports, Volume 60, 2019, Pages 46-48. https://doi.org/10.1016/j.ijscr.2019.05.061.

[6]. De Potter P, Cools H, Depraetere K, Mels G, Debevere P, De Roo J, Huszka C, Colaert D, Mannens E, Van de Walle R. Semantic patient information aggregation and medicinal decision support. Comput Methods Programs Biomed. 2012 Nov;108(2):724–35.

[7]. Dellamonica P, le Fichoux Y, Coussement A, Serres JJ. Diagnostic des calcifications parasitaires (à l'exclusion des calcifications fongiques et des protozoaires) [Parasitic calcifications: elements for diagnostic (author's transl)]. J Radiol Electrol Med Nucl. 1977;58(2):125-133.

[8]. Dev B, Vakamudi UB, Thambidurai L, Joseph LD, Srinivasan J. Imaging and Pathological Correlation in Spectrum of Fibrocystic Breast Disease and its Mimics–our Experience. Archives of Breast Cancer. 2022 Sep 26:465-73.

[9]. Egshatyan L.V., Mokrysheva N.G. ECTOPIC CALCIFICATION IN CHRONIC KIDNEY DISEASE. PART 1. CLASSIFICATION AND PATHOGENESIS. Nephrology (Saint-Petersburg). 2017;21(4):30-39. (In Russ.)

[10]. Faria de LL, Babler F, Ferreira LC, de Noronha Junior OA, Marsolla FL, Ferreira DL. Soft tissue calcifications: a pictorial essay. Radiol Bras. 2020 Sep-Oct;53(5):337-344.

[11]. Giachelli CM. Ectopic calcification: gathering hard facts about soft tissue mineralization. Am J Pathol. 1999 Mar;154(3):671-5.

[12]. Golbeck J, Fragoso G, Hartel F, Hendler J, Oberthaler J, Parsia B. The National Cancer Institute's Thésaurus and Ontology. Journal of Web Semantics. 2003 Dec;1(1):75–80.

[13]. Gosling S, Scott R, Greenwood C, et al. Calcification Microstructure Reflects Breast Tissue Microenvironment. J Mammary Gland Biol Neoplasia. 2019;24(4):333-342. doi:10.1007/s10911-019-09441-3.

[14]. Guerlain J, Perie S, Lefevre M, et al. Localization and characterization of thyroid microcalcifications: A histopathological study. PLoS One. 2019;14(10):e0224138. Published 2019 Oct 24. doi:10.1371/journal.pone.0224138.

[15]. Guerlain J, Perie S, Lefevre M, Perez J, Vandermeersch S, Jouanneau C, Huguet L, Frochot V, Letavernier E, Weil R, Rouziere S, Bazin D, Daudon M, Haymann JP. Localization and characterization of thyroid microcalcifications: A histopathological study. PLoS One. 2019 Oct 24;14(10):e0224138. doi: 10.1371/journal.pone.0224138.

[16]. Infectious Disease Ontology (2020). https://bioportal.bioontology.org/ontologies/IDO. Accessed 12 Nov 2023

[17]. Isern D, Sánchez David, Moreno A. Ontology-driven execution of clinical guidelines. Comput Methods Programs Biomed. 2012 Aug;107(2):122–39.

[18]. Kuhn M, Letunic I, Jensen LJ, et al. The SIDER database of drugs and side effects. Nucleic Acids Res 2016;44(D1):D1075–9.

[19]. Lanzer P, Boehm M, Sorribas V et al. Medial vascular calcification revisited: review and perspectives. European Heart Journal 2014; 35: 1515-1525

[20]. Li Y, Cao J, Zhou Y, Mao F, Shen S, Sun Q. Mammographic casting-type calcification is an independent prognostic factor in invasive breast cancer. Sci Rep. 2019 Jul 22;9(1):10544.

[21]. National library of medicine(2023) Medical subject headings (MeSH) Thesaurus. https://www.nlm.nih.gov/mesh/meshhome.html. Accessed on 18 Dec 2023

[22]. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2018 Jan 4;46(D1):D8-D13. doi: 10.1093/nar/gkx1095.

[23]. Oksanchuk E.A., Meskih E.V., Kolesnik A.U., Nudnov N.V. Breast Calcifications: Etiology, Classification, Prognosis. Medical Visualization. 2016;(5):120-127. (In Russ.).

[24]. Ontology browser(2024) https://rgd.mcw.edu/rgdweb/ontology/view.html?acc_id=DOID:182. Accessed on 9 Feb 2024

[25]. Poux S, Arighi CN, Magrane M, Bateman A, Wei C-H, Zhiyong L, Boutet E, Bye-A-Jee H, Famiglietti ML, Roechert B, The UniProt Consortium. On expert curation and scalability: UniProtKB/Swiss-Prot as a case study. Bioinformatics. 2017;33(21):3454–60.

[26]. Poux S, Magrane M, Arighi CN, Bridge A, O'Donovan C, Laiho K, The UniProt Consortium. Expert curation in UniProtKB: a case study on dealing with conflicting and erroneous data. Database. 2014;2014:bau016.

[27]. Proudfoot D. Calcium Signaling and Tissue Calcification. Cold Spring Harb Perspect Biol. 2019 Oct 1;11(10):a035303.

[28]. Radiology assistant (2023) https://radiologyassistant.nl/breast/calcifications/differential-of-breast-calcifications. Accessed on 18 Dec 2023

[29]. Radiology Assistant (2024) https://radiologyassistant.nl/breast/calcifications/differential-of-breast-calcifications. Accessed on 27 Dec 2023.

[30]. Radiopaedia (2024) https://radiopaedia.org/articles/amorphous-calcifications-breast?lang=us. Accessed on 9 Jan 2024

[31]. Richesson RL, Nadkarni P. Data standards for clinical research data collection forms: current status and challenges. J Am Med Inform Assoc. 2011;18(3):341–346.

[32]. Rosse C, Mejino JL. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. J Biomed Inform. 2003 Dec;36(6):478–500. DOI: 10.1016/j.jbi.2003.11.007

[33]. Sandun I, Sumathipala S, , Ganegoda GU. Self-evolving Disease Ontology for Medical Domain Based on Web. IJFIS 2017;17:307-314

[34]. Schulz S, Cornet R, Spackman K. Consolidating SNOMED CT's ontological commitment. Applied Ontology. 2011 Jan;6(1):1–11.

[35]. Senba M, Kawai K, Mori N. Pathogenesis of Metastatic Calcification and Acute Pancreatitis in Adult T-Cell Leukemia under Hypercalcemic State. Leuk Res Treatment. 2012;2012:128617.

[36]. Shi X, Gao J, Lv Q, Cai H, Wang F, Ye R, Liu X. Calcification in Atherosclerotic Plaque Vulnerability: Friend or Foe? Front Physiol. 2020 Feb 5;11:56.

[37]. Suh JW, Yun B. Breast Arterial Calcification: A Potential Surrogate Marker for Cardiovascular Disease. J Cardiovasc Imaging. 2018 Sep;26(3):125-134. doi: 10.4250/jcvi.2018.26.e20.

[38]. Sumathipala S, Yamada K, Unehara M, Suzuki I. Protein entity name recognition using orthographic, morphological and proteinhood features. Journal of Advanced Computational Intelligence and Intelligent Informatics. 2015; 19, 843-851.

[39]. Suresh SS, Raniga S, Shanmugam V, George M, Zaki H. Carpal tunnel syndrome due to hydroxyapatite crystal deposition disease. J Hand Microsurg. 2013 Dec;5(2):96-9. doi: 10.1007/s12593-012-0076-9.

[40]. Tandon R, Agaral S, Goyal P. Sequential convolutional neural network for automatic breast cancer image classification using histopathological images. methods. 2020;7(10):10-31838.

[41]. The Human Phenotype Ontology (2024). https://hpo.jax.org/app/. Accessed on 27 Jan 2024

[42]. Translational Medicine Ontology (2016). https://bioportal.bioontology.org/ontologies/TMO. Accessed on 18 Jan 2024

[43]. Tsuruoka, Y, and Tsujii, JI 2003. Boosting precision and recall of dictionary-based protein name recognition., Proceedings of the ACL 2003 Workshop on Natural Language Processing in Biomedicine, Sapporo, Japan, Array, pp.41-48.

[44]. Wang P, Hu Y, Bai S, Zou S. Matching Biomedical Ontologies: Construction of Matching Clues and Systematic Evaluation of Different Combinations of Matchers. JMIR Med Inform. 2021 Aug 19;9(8):e28212. DOI: 10.2196/28212

[45]. Whirl-Carrillo M, McDonagh EM, Hebert JM, et al. Pharmacogenomics knowledge for personalized medicine. Clin Pharmacol Ther 2012;92(4):414–7.

[46]. Yang Z, Lin H, Li Y. Exploiting the performance of dictionary-based bio-entity name recognition in biomedical literature. Computational Biology and Chemistry. 2008; 32, 287-291.

[47]. Yongjun Zhu, Olivier Elemento, Jyotishman Pathak, Fei Wang. Drug knowledge bases and their applications in biomedical informatics research, Briefings in Bioinformatics, Volume 20, Issue 4, July 2019, Pages 1308–1321.

[48]. Yu MH, Kim YJ, Park HS, Jung SI, Jeon HJ. Imaging Patterns of Intratumoral Calcification in the Abdominopelvic Cavity. Korean J Radiol. 2017 Mar-Apr;18(2):323-335.