MEDINFO 2021: One World, One Health – Global Partnership for Digital Innovation
P. Otero et al. (Eds.)

doi:10.3233/SHTI220084

A Knowledge Graph of Mechanistic Associations Between COVID-19, Diabetes Mellitus, and Chronic Kidney Disease

Michael Barrett[a], Syed Sibte Raza Abidi[a], Ali Daowd[a], Samina Abidi[ab]

[a] *NICHE Research Group, Faculty of Computer Science, Dalhousie University, Halifax, Nova Scotia, Canada*
[b] *Department of Community Health and Epidemiology, Faculty of Medicine, Halifax, Nova Scotia, Canada*

Abstract

We present an automated knowledge synthesis and discovery framework to analyze published literature to identify and represent underlying mechanistic associations that aggravate chronic conditions due to COVID-19. Our literature-based discovery approach integrates text mining, knowledge graphs and medical ontologies to discover hidden and previously unknown pathophysiologic relations, dispersed across multiple public literature databases, between COVID-19 and chronic disease mechanisms. We applied our approach to discover mechanistic associations between COVID-19 and chronic conditions—i.e. diabetes mellitus and chronic kidney disease—to understand the long-term impact of COVID-19 on patients with chronic diseases. We found several gene-disease associations that could help identify mechanisms driving poor outcomes for COVID-19 patients with underlying conditions.

Keywords:

Literature Based Discovery, COVID-19, Diabetes Mellitus.

Introduction

The SARS-CoV-2 virus (causative agent of COVID-19) has the ability to target multiple organs while activating, in some cases, an intense and systemic immune response. The virus uses the ACE2 receptor to gain entry into cells, which is expressed in the lungs, liver, endocrine pancreas, kidney, endothelium and heart [1], leading to acute multiorgan injuries and death [2], caused by some combination of direct viral involvement, systemic proinflammatory immune response, or systemic hypoxia and coagulopathy [3]. SARS-CoV-2 infections are more likely to progress to severe illness in patients with Diabetes Mellitus (DM) and Chronic Kidney Disease (CKD) [4]. At present, the basic mechanisms, pathophysiological pathways and processes by which COVID-19 *exacerbates* or is *exacerbated by* DM and CKD are not well known [5], despite the rapid accumulation of published evidence on COVID-19. Researchers are actively seeking tools to discover from published evidence how mechanisms involving virus-host interactions, gene-environment networks, and metabolic and immune signaling—referred to as *mechanistic associations*—contribute to increased risk for DM and CKD patients [6].

Related Work

Discovering new knowledge to understand the mechanistic associations between COVID-19 and DM, CKD or both is challenging as the rapidly emerging evidence is distributed across a large number of publications that contain many explicit and implicit causal relationships between diseases, comorbidities and molecular entities (e.g. genes, proteins and metabolites). Recent attempts to integrate COVID-19 literature have resulted in web-based tools like the COVID-19 Knowledge Graph [7] that stores causal mechanistic associations that were manually extracted from the literature and presented as a visual semantic network. The advantage of semantic graphs is that they represent knowledge in a computable form such that it can be queried to represent complex associations (e.g. disease mechanisms), which is infeasible to do in traditional (relational) databases. In a previous work [8], an analytical method leveraging text mining methods was used to find biochemical associations, dispersed across multiple knowledge sources, by integrating manually curated associations in a graph. This allowed the authors to uncover biomarkers using background knowledge to supplement relations found in text. Moreover, advanced methods can find complex molecular associations involving a series of related concepts, discovering unnoticed links between disease mechanisms. Recently, Baek et al. [9] captured associations using semantic relatedness and graph-based methods, allowing them to uncover novel biological pathways.

Related works [8], [9] seek to automatically discover useful relations which are subsequently validated through evidence-based means. This typically involves the use of text mining tools and pattern analysis to identify biologically meaningful associations hidden in large bodies of literature, using pre-specified ranking criteria (e.g. frequency- or graph-based metrics) to focus on important discoveries. Inspired by recent studies, our goal is to extend the proposed methods to improve literature based discovery of mechanistic associations by discovering plausible disease mechanisms that can help understand the pathobiological mechansisms underlying disease impacts.

This paper presents a novel COVID-19 knowledge synthesis and discovery approach that integrates (a) Literature-Based Discovery (LBD) [10] to automatically extract both known and new mechanistic associations from published medical literature, such as PubMed; (b) Medical ontologies to further augment the mechanistic associations; and (c) Knowledge Graph (KG) to represent the mechanistic associations. This approach was subsequently implemented as a LBD framework [10]. We applied our framework to the rapidly evolving COVID-19 literature to answer two queries—(1) what are the drivers of COVID-19 progression in the context of molecular and physiological perturbations that are associated with DM and CKD, and (2) how might these mechanistic associations be implicated in previously hypothesized mechanisms?

Methods

We developed a LBD framework (Figure 1) that combines text mining and ontologies for knowledge discovery from medical literature. The discovered mechanistic associations between between COVID-19, DM and CKD are represented as a KG. We have developed pattern analysis methods to traverse the KG in response to user queries seeking answers about the impact of COVID-19 patient with DM and CKD.

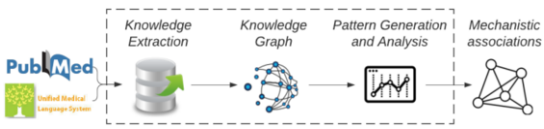

Figure 1– LBD framework

Knowledge Discovery from Medical Literature

We analyzed research articles related to COVID-19 and chronic diseases in PubMed (2020 – onwards). We used a NLP tool called SemRep to retrieve article identifiers (i.e. PubMed IDs) that were used to index the Semantic MEDLINE Database (SemMedDB), a repository of semantic associations extracted from PubMed abstracts [11]. Given textual abstracts as input, SemRep reads each sentence and infers logical relationships—called *semantic associations*—between concepts of interest. For instance, given the sentence *"The capacity for autophagy in both podocytes and renal tubular cells is markedly impaired in type 2 diabetes"*, it provided:

1. Tubular Cells (subject)–LOCATION_OF (relation)→ Autophagy (object)
2. Podocytes (subject)–LOCATION_OF (relation)→ Autophagy (object).

Each association consists of a standardized subject and object mapped to Unified Medical Language System (UMLS) semantic types (e.g. *cytokine production* is mapped to *Organism Function*) and a high-level relation type. SemRep extracts assertions based on relations in comorbidity (e.g. COEXISTS_WITH), substance interactions (e.g. INTERACTS_WITH) and physiologic disturbances (e.g. DISRUPTS). As such, the user is provided with concise statements about biological relationships that are used to identify interesting associations between concepts, typically by navigating with prior background knowledge.

A novel aspect of our approach is the integration of external knowledge sources (ontologies) and public databases to further augment the semantic associations to draw out interesting relations regarding gene function. The use of ontologies allows us to extract implicit associations that remain hidden from typical LBD approaches. Our method augments associations that are simplistic or uninformative with regards to biomedically important relationships (e.g. Diabetes – AFFECTS → Immune Response) by automatically providing alternative (i.e. related) concepts. For instance, the concept 'Immune Response' will be extended to include 'Complement Activation'. This step involved semantic integration using the UMLS MRREL dataset that comprises associations between biomedical vocabularies (i.e. ontologies). Gene function and pathway concepts were linked to semantic associations where applicable by expanding physiologic concepts related to COVID-19, DM, and CKD, using hierarchical (e.g. narrower) and associative relations from the UMLS. Finally, gene and gene function concepts were used as input to retrieve ontology annotations from a public database. As a result, semantic associations were enriched with biomedical knowledge specific to DM, CKD, and COVID-19 to produce a set of highly interrelated concepts.

Representing Associations as a Knowledge Graph

To capture relations between concepts to represent complex associations, the semantic associations were represented as an interactive KG using a Neo4j graph database—the KG represents each subject and object concept as a node and the predicate (i.e. semantic association/relation between the concepts) as a directed link between nodes. Each concept node was annotated with its name, UMLS identifier, semantic type, and frequency of occurrence. Relations were assigned a type, frequency of occurrence, and unique identifier (article PMID or database reference) to establish provenance. Finally, concepts were assigned to one of five high-level UMLS semantic groups (*anatomy*, *chemicals*, *disorders*, *genes and molecular sequences*, or *physiology*) to filter out uninteresting or irrelevant semantic types (e.g. *Laboratory Procedure*). For example, proteins and metabolites were assigned to chemicals and diseases and comorbidities to disorders.

Discovering Mechanistic Associations from the KG

To extract mechanistic associations from the KG, we developed a search strategy using Neo4j's query language Cypher whereby we can specify both direct (concept A → B) or indirect (A → … → B) associations as multi-node patterns—called *discovery patterns*—that comprise substance interactions, physiologic disturbances, and disease-disease relations (e.g. comorbidity). Direct associations were specified by searching for a given pair of concepts with specific relations between them (e.g. a gene and its associated physiologic function). We combined direct associations by defining logical sequences of relations to find genes indirectly (2 or more concepts away) associated with COVID-19, DM or CKD, which refers to molecular mechanisms (e.g. gene-disease associations) that may be implicated in a disease.

Table 1 presents the top 5 discovered semantic associations.

Table 1– Patterns of semantic associations linking genes, chemicals, physiology, and disorders

Node 1	Relation 1	Node 2	Relation 2	Node 3	Relation 3	Node 4
Complement 3 gene	*Has annotation*	Alternative Complement Pathway	*Is a*	Immunity, Innate	*Associated with*	COVID-19
NF-kappa B	*Stimulates*	SIRT1 gene	*Augments*	Oxidative Stress	*Coexists with*	COVID-19
NLRP3 gene	*Has annotation*	Negative regulation of inflammatory response	*Associated with*	COVID-19	–	–
Interleukin-1, beta	*Stimulates*	FGF23 gene	*Disrupts*	Vitamin D metabolic process	–	–
Iron	*Interacts with*	FGF23 gene	*Interacts with*	Vitamin D	*Augments*	Osteoblast differentiation

By combining associations we are able to identify associations that have not yet been captured and which could be inferred based on the given relations. An example of an indirect (mechanistic) association being discovered is as follows: (1) Gene A *augments* Biologic Function; (2) Biologic Function *causes* Disease B; and (3) Gene A *may affect* Disease B.

To identify interesting and important discovery patterns we developed a ranking mechanism that compares (i) indirect association and (ii) graph-theoretic measures, namely Linking Term Count (LTC) [12] and PageRank [13], respectively. LTC considers the number of intermediate concepts when A and B are 2 concepts away to assess whether A and B are strongly correlated. PageRank is a measure of each node's connectivity in the network, which tells if a given concept and its directly related concepts are highly important. We used the average scores of each pattern to rank the associations as a criterion for discovery of mechanistic associations, which was calculated as follows: (1) Count intermediate concepts (x) between A and B (i.e. A – x – B); (2) Calculate sum of PageRank of all nodes in the pattern and divide by the number of nodes; and (3) Compute average score using the two metrics.

Results

We found 11,403 papers on COVID-19, DM or CKD, of which 115 were on COVID-19 related to DM or CKD. Using semantic filtering and prior background knowledge to focus on highly relevant concepts (e.g. genes, physiologic functions, disorders), we retrieved 2,089 unique associations from these articles in SemMedDB and represented them in a KG. The result was a complex network of comorbid conditions with molecular, physiologic, and pathologic levels. The two ranking techniques initially gave contrasting scores, which was expected since the two metrics are calculated in different ways. To reduce noise, we only considered patterns where PageRank and LTC scores deviated by less than a factor of 10. It was noted that LTC rankings were more heterogeneous than PageRank with regards to semantic associations retrieved from external knowledge sources (ontologies, annotations) while PageRank tended to favour patterns containing well-connected concepts from the literature (e.g. COVID-19). We identified 26 patterns in total, including gene-physiology and gene-disease (indirect) associations. The five top-ranking discovery patterns that are regarded as candidate mechanistic associations are shown in Table 1. The ranking scores of these patterns (in order of PageRank score as in Table 1) are shown in Table 2, and potential discoveries are discussed in detail below.

Table 2– Ranking scores of discovery patterns

Position	LTC	PageRank	Average Score
1	0	6.5	3.92
2	8	6.45	7.23
3	4	0.51	2.25
4	1	0.18	0.59
5	0	0.17	0.09

In Table 1, the first relation is referring to a biological role of a given gene, in this case Complement 3 (*C3*). As per our approach, this relation was annotated using the Gene Ontology (GO) [14], which linked to evidence of a gene's function(s) to provide an implicit mechanistic association through related concepts in the literature. While the association between *C3* and innate immunity is accurate, the role of the alternative complement pathway in COVID-19 is unclear, which could be due to the overall (gene-disease) association being inconsistent or poorly correlated (as per its low LTC score). Through extension with an ontology relation (Relation 2), the GO annotation was integrated with an association extracted from the literature (i.e. SemMedDB) shown on the far right (Relation 3). The SemMedDB article linking innate immunity to COVID-19 was not focused on the complement system other than to briefly discuss the role of abnormal coagulation mediated by complement activation in response to infection [15]. The presence of *C3* is explained by a relation (not shown) from a different study showing that complement activation is implicated in oxidative stress and early renal decline in mice with Diabetic Nephropathy (DN) [16]. The extended relation involving *C3* thusly provides a similar association to the one mined from text to integrate associations that are described in related contexts (i.e. pathways of innate immunity). The inferred relation would be that *C3* may affect COVID-19 by interacting with the alternative complement pathway.

The second pattern in Table 1 comprises a series of relations that were found in SemMedDB. The accuracy of these relations varies since they were extracted from published articles by text mining without prior validation. The first relation is inaccurate since NF-kappa B was not found to stimulate sirtuin 1 (*SIRT1*). However, there is a correlation between NF-kappa B and oxidative stress and *SIRT1* may regulate this pathway in adipose tissue [17], suggesting that *SIRT1* does not augment oxidative stress and may disrupt it instead. This hypothesis is relevant to DM patients since the preceding study demonstrated a link between obesity and decreased *SIRT1* expression, which was thought to contribute to abnormal adipose tissue inflammation. Finally, abnormal oxidative stress and inflammation may cause patients with underlying chronic conditions to experience worse COVID-19 outcomes [18]. Thus, the inferred relation would be that NF-kappa B may affect COVID-19 by inducing oxidative stress, which could be regulated by *SIRT1*.

The third pattern in Table 1 shows a gene-disease association between *NLRP3* and COVID-19. The first relation originated from GO and referred to a published study showing that *NLRP3* mediates cytokine signaling in primary immune cells [19]. The second relation was retrieved from SemMedDB and linked to an article showing that patients with severe COVID-19 expressed stable levels of pro- and anti-inflammatory cytokines, which was hypothesized to protect these patients from inflammatory tissue damage [20]. The preceding article did not make reference to *NLRP3* activity and it is therefore unclear whether *NLRP3* participates in the hypothesized mechanism. The inferred relation supposes that *NLRP3* fails to downregulate inflammation following exposure to COVID-19, potentially leading to long-term impacts on chronic conditions, though preliminary evidence suggests this is not the case [20].

The final two patterns in Table 1 show gene-physiology associations that may be relevant to COVID-19. Interleukin-1 beta, a systemic regulator of inflammation, may induce *FGF23* production in early CKD in mice, which could be followed by disruptions to vitamin D metabolism [21]. This hypothesis may be important for COVID-19, DM and CKD since vitamin D has anti-inflammatory effects which could be disrupted by vitamin D deficiency in these disorders. The inferred relation for the fourth pattern would be that interleukin-1 beta disrupts vitamin D processing through *FGF23*, potentially leading to chronic inflammation. The fifth pattern proposes that iron has an indirect effect on osteoblast differentiation through *FGF23*, which is speculative since the paper linking vitamin D to osteoblast differentiation did not mention iron.

Figure 2 shows a segment of the KG that was generated to capture complex associations; all associations shown in the

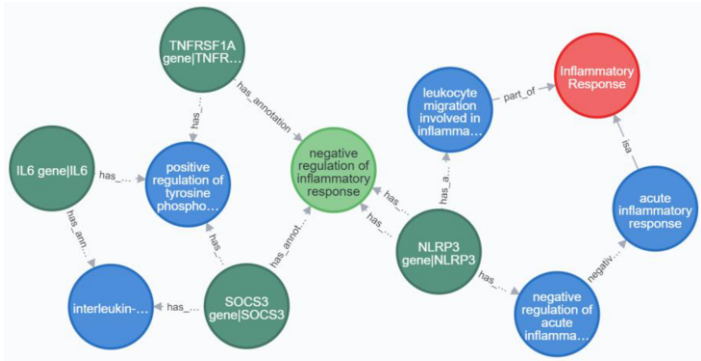

Figure 2– Segment of the KG. Dark green nodes: genes. Light green/Blue nodes: physiology. Red node: disorders

figure originated from GO or the UMLS. The cluster of nodes on the left depicts genes involved in cytokine-mediated signaling pathways, where the associations with GO terms (shown in blue) in this cluster were automatically annotated. The central blue node in this cluster is a physiologic component of the JAK-STAT pathway. The *TNFRSF1A* gene (encodes *TNFR1*) was implicated in an unrelated inflammatory disorder [22] by its GO association. *SOCS3* is a negative regulator of cytokine signaling that may be relevant to DM or CKD, though its association with COVID-19 is unclear at the moment. The association between *NLRP3* and *leukocyte migration involved in inflammatory response* (as shown in Figure 2) was assigned electronically and lacks evidence. Finally, *NLRP3* was indirectly associated with the abovementioned genes through negative regulation of proinflammatory cytokines. We continue to explore visualization strategies in ongoing work. Figure 3 shows a segment of the KG involving *IL6, CD14*, and Type 2 DM, which is being investigated and could have implications for underlying mechanisms of COVID-19.

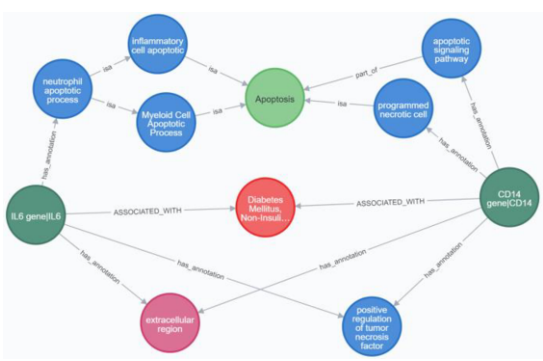

Figure 3– Segment of the KG showing ongoing work.

Discussion

Our LBD approach brought together concepts from distinct knowledge sources to generate mechanistic associations. The integration of semantic associations from an existing resource (SemMedDB) with external knowledge allowed us to uncover mechanistic associations including five gene-disease associations. By comparison, a recent LBD method brought together three biochemical relations to represent a novel biological pathway which was validated in clinical tests [9]. Previous LBD studies only consider indirect links as two concepts away, such as the work to discover biomarkers for migraine [8], using a single pattern ranking technique. In our work, we have generated patterns (i.e. testable indirect associations) to link distant literature sources and thus increase the likelihood of discovering hidden or unknown associations, whilst also using multiple pattern-ranking methods. Our approach found several associations that may be relevant to mechanisms underlying COVID-19, though the suggested associations were often tentative or inconsistent. For instance, *C3* was thought to play a role in abnormal complement activation in DN [16], but this pathway was not immediately relevant to the association between COVID-19 and innate immunity [15]. Interestingly, *SIRT1* deficiency was found to increase levels of *C3* in adipose tissue in mice [17]; however, it is unclear whether altered levels of *C3* exist in COVID-19 patients with underlying conditions. Further, the association between *NLRP3* and negative regulation of inflammation was not explicitly related to COVID-19 as it applied to a group of patients with unrelated disorders [19].

Taken together, these findings underscore the need for careful selection of gene and physiologic relations to identify potential disease pathways. We were able to identify plausible links between molecular and pathophysiologic concepts that were enriched with contextual information (i.e. the nature of those relationships) from SemMedDB. However, the relation extraction task was prone to error when assessing the relations between chemicals (NF-kappa B, *SIRT1*) and disorders (Oxidative Stress). Moreover, there was one issue related to the accuracy of an electronically assigned relation in GO. Future work will involve better curation strategies to focus on high quality relations where possible. This could involve further iterations of knowledge extraction to find interesting concepts in SemMedDB and GO, filtering based on relations in the literature, and statistical tests to narrow down the list of genes to a highly relevant subset.

Different ranking methods prioritized unique patterns of associations. Including ontologies may create distinct patterns that are made apparent by using different ranking techniques, which is supported by our findings since PageRank tended to focus on well-connected literature concepts while LTC prioritized associations completely differently with regards to interesting associations from GO. As such, it may be beneficial to have a variety of metrics to choose from in certain scenarios. However, it could become difficult to assess rankings objectively should these experiments be repeated with a larger dataset. This will be addressed by standardizing rankings to produce a more robust prediction score. Finally, future experiments will investigate whether different ranking techniques can be combined to single out interesting patterns.

The majority of indirect associations were found from more than two concepts away, and interacting with the KG identified interesting and complex mechanisms. We noticed that patterns involving only three concepts were unable to capture complex associations such as the ones shown in Figure 2 and Figure 3. Other complex associations could remain undetected (e.g. genes participating in similar physiologic functions). On the other hand, it appears reasonable to investigate both simple and complex patterns to focus on interesting findings. For example, the link between *NLRP3* and COVID-19 in Table 1 was not explained in related literature, which suggests this association may be relevant to different disorders. Therefore, focusing on simple patterns to establish relevance is expected to help narrow down the list of mechanistic associations to identify promising complex patterns.

There are several limitations to this work. Certain studies of COVID-19 patients with DM or CKD involved small sample sizes, which could limit the applicability of these findings to a wider population. Further, extending semantic associations to cover related concepts relies on multiple domain knowledge sources (ontologies, annotations) that may not be available in other research contexts. Creating patterns is a time-consuming process based on relations that are known ahead of time, which limits the generalizability of our methods. Finally, interacting with a KG to construct interesting visualizations is manually intensive and may not be feasible with larger datasets.

Conclusion

This paper presents a LBD framework that offers a novel integration of published literature with biomedical knowledge to synthesize published evidence to discover new knowledge about the interplay between COVID-19 and chronic conditions. We integrated relations that were dispersed across multiple public databases, using pattern analysis and ranking to uncover plausible disease mechanisms. We found several gene candidates that appear to be interrelated through disease states and physiologic processes that are relevant to COVID-19 and DM or CKD. More research is needed to understand how genetic and immunologic factors contribute to worse COVID-19 outcomes for patients with DM or CKD.

Our knowledge discovery approach is disease agnostic and can be applied to investigate mechanistic associations from the literature for any disease. We conclude that further investigation is needed to identify meaningful extensions of semantic associations and to assess the relative performance of LTC and PageRank as our initial results suggest that both are able to focus the user's attention on highly relevant patterns.

References

[1] M. Iwai, and M. Horiuchi, Devil and angel in the renin-angiotensin system: ACE-angiotensin II-AT1 receptor axis vs. ACE2-angiotensin-(1-7)-Mas receptor axis, *Hypertens. Res.* **32** (2009) 533–536.

[2] N. Potere, et al, Acute complications and mortality in hospitalized patients with coronavirus disease 2019: A systematic review and meta-analysis, *Crit. Care*. **24** (2020) 389.

[3] T. Wu, et al, Multi-organ Dysfunction in Patients with COVID-19: A Systematic Review and Meta-analysis, *Aging Dis.* **11** (2020) 874-894.

[4] A.K. Singh, et al, Prevalence of co-morbidities and their association with mortality in patients with COVID-19: A systematic review and meta-analysis, *Diabetes, Obes. Metab.* **22** (2020) 1915-1924.

[5] R. Gupta, et al, Diabetes and COVID-19: evidence, current status and unanswered research questions, *Eur. J. Clin. Nutr.* (2020) 864-870.

[6] M. Apicella, et al, COVID-19 in people with diabetes: understanding the reasons for worse outcomes, *Lancet Diabetes Endocrinol.* **8** (2020) 782–792.

[7] D. Domingo-Fernandez, et al, COVID-19 Knowledge Graph: a computable, multi-modal, cause-and-effect knowledge model of COVID-19 pathophysiology, *BioRxiv*. (2020).

[8] W.J. Vlietstra, et al, Automated extraction of potential migraine biomarkers using a semantic graph, *J. Biomed. Inform.* **71** (2017) 178–189.

[9] [9] S.H. Baek, et al, Enriching plausible new hypothesis generation in PubMed, *PLoS One*. **12** (2017).

[10] S. Henry, et al, Literature Based Discovery: Models, methods, and trends, *J. Biomed. Inform.* **74** (2017) 20–32.

[11] H. Kilicoglu, et al, SemMedDB: A PubMed-scale repository of biomedical semantic predications, *Bioinformatics*. **28** (2012) 3158–3160.

[12] D.R. Swanson, and N.R. Smalheiser, An interactive system for finding complementary literatures: A stimulus to scientific discovery, *Artif. Intell.* **91** (1997) 183–203.

[13] L. Page, and S. Brin, The anatomy of a large-scale hypertextual Web search engine, *Com. Net*. **30** (1998) 107–117.

[14] S. Carbon, et al, The Gene Ontology resource: Enriching a GOld mine, *Nucleic Acids Res.* **49** (2021) D325–D334.

[15] P.R. Criado, et al, Lessons from dermatology about inflammatory responses in Covid-19, *Rev. Med. Virol.* **30** (2020).

[16] M. Morigi, et al, C3a receptor blockade protects podocytes from injury in diabetic nephropathy, *JCI Insight*. **5** (2020).

[17] M.P. Gillum, et al, SirT1 regulates adipose tissue inflammation, *Diabetes*. **60** (2011) 3235–3245.

[18] C. Lammi, and A. Arnoldi, Food-derived antioxidants and COVID-19, *J. Food Biochem.* **45** (2021) 1–6.

[19] Aksentijevich, et al, De novo CIAS1 mutations, cytokine activation, and evidence for genetic heterogeneity in patients with neonatal-onset multisystem inflammatory disease (NOMID), *Arthritis Rheum.* **46** (2002) 3340–3348.

[20] Q. Notz, et al, Pro- and Anti-Inflammatory Responses in Severe COVID-19-Induced ARDS—An Observational Pilot Study, *Front. Immunol.* **11** (2020).

[21] Q. McKnight, et al, IL-1β Drives Production of FGF-23 at the Onset of Chronic Kidney Disease in Mice, *J. Bone Miner. Res.* **35** (2020) 1352–1362.

[22] E. Aganna, et al, Heterogeneity among patients with tumor necrosis factor receptor-associated periodic syndrome phenotypes, *Arthritis Rheum.* **48** (2003) 2632–2644.

Address for correspondence

Syed Abidi, Faculty of Computer Science, Dalhousie University, Halifax, Canada. Email: ssrabidi@dal.ca.