



OPEN ACCESS

EDITED BY

Ritthideach Yorsaeng,
Chulalongkorn University, Thailand

REVIEWED BY

Sunil Gairola,
Serum Institute of India, India
Pasquale Bufano,
National Research Council (CNR), Italy

*CORRESPONDENCE

Meredith Leston

✉ meredith.leston@phc.ox.ac.uk

RECEIVED 09 September 2024

ACCEPTED 14 November 2024

PUBLISHED 04 December 2024

CITATION

Leston M, Kar D, Forbes A, Jamie G, Wimalaratna R, Jiwani G, Ordóñez-Mena JM, Stewart DE, Whitaker H, Joy M, Lee LYW, Hobbs FDR and de Lusignan S (2024) Developing a digital phenotype to subdivide adult immunosuppressed COVID-19 outcomes within the English Primary Care Sentinel Network. *Front. Immunol.* 15:1491565. doi: 10.3389/fimmu.2024.1491565

COPYRIGHT

© 2024 Leston, Kar, Forbes, Jamie, Wimalaratna, Jiwani, Ordóñez-Mena, Stewart, Whitaker, Joy, Lee, Hobbs and de Lusignan. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Developing a digital phenotype to subdivide adult immunosuppressed COVID-19 outcomes within the English Primary Care Sentinel Network

Meredith Leston^{1*}, Debasish Kar¹, Anna Forbes¹, Gavin Jamie¹, Rashmi Wimalaratna¹, Gunjan Jiwani¹, José M. Ordóñez-Mena¹, Daniel E. Stewart^{2,3}, Heather Whitaker², Mark Joy¹, Lennard Y. W. Lee⁴, F. D. Richard Hobbs¹ and Simon de Lusignan¹

¹Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom, ²Immunisation and Vaccine Preventable Diseases Division, UK Health Security Agency, London, United Kingdom, ³UK Field Epidemiology Training Programme (UK-FETP), UK Health Security Agency, London, United Kingdom, ⁴Department of Oncology, University of Oxford, Oxford, United Kingdom

Background: Adults classified as immunosuppressed have been disproportionately affected by the COVID-19 pandemic. Compared to the immunocompetent, certain patients are at increased risk of suboptimal vaccine response and adverse health outcomes if infected. However, there has been insufficient work to pinpoint where these risks concentrate within the immunosuppressed spectrum; surveillance efforts typically treat the immunosuppressed as a single entity, leading to wide confidence intervals. A clinically meaningful and computerised medical record (CMR) compatible method to subdivide immunosuppressed COVID-19 data is urgently needed.

Methods: We conducted a rapid scoping review into COVID-19 mortality across UK immunosuppressed categories to assess if differential mortality risk was a viable means of subdivision. We converted the risk hierarchy that surfaced into a pilot digital phenotype—a valueset and series of ontological rules ready to extract immunosuppressed patients from CMR data and stratify outcomes of interest in COVID-19 surveillance dataflows.

Results: The rapid scoping review returned COVID-19 mortality data for all immunosuppressed subgroups assessed and revealed significant heterogeneity across the spectrum. There was a clear distinction between heightened COVID-19 mortality in haematological malignancy and transplant patients and mortality that approached the immunocompetent baseline amongst cancer therapy recipients, autoimmune patients, and those with HIV. This process, complemented by expert clinical input, informed the curation of the five-part digital phenotype now ready for testing in real-world data; its ontological rules will enable mutually exclusive, hierarchical extraction with nuanced time and treatment conditions. Unique categorisations have been introduced, including

'Bone Marrow Compromised' and those dedicated to differentiating prescriptions related and unrelated to cancer. Codification was supported by existing reference sets of medical codes; absent or redundant codes had to be resolved manually.

Discussion: Although this work is in its earliest phases, the development process we report has been highly informative. Systematic review, clinical consensus building, and implementation studies will test the validity of our results and address criticisms of the rapid scoping exercise they are predicated on.

Conclusion: Comprehensive testing for COVID-19 has differentiated mortality risks across the immunosuppressed spectrum. This risk hierarchy has been codified into a digital phenotype for differentiated COVID-19 surveillance; this marks a step towards the needs-based management of these patients that is urgently required.

KEYWORDS

digital health, vaccine, immunosuppressed, disease surveillance, surveillance, CMR

Introduction

Despite it being well-known that the immunosuppressed population experience worse respiratory infection outcomes, large administrative or surveillance datasets rarely collect data on how these might vary between condition types (1–4). Reports are either aggregated, handling immunosuppressed data as a singular entity (5), or only return differentiated data for dominant subgroups, such as cancers (6). This erases wider sub-trends that may be of clinical interest and undermines evidence-based prioritisation of medical resources including vaccines and their passive alternatives (7).

The Oxford-Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) is well-positioned to respond to recent appeals (8) made by its partners, the United Kingdom Health and Security Agency (UKHSA) to improve the granularity of immunosuppressed COVID-19 surveillance. The trusted research environment it utilises, the Oxford-RCGP Clinical Informatics Digital Hub (ORCHID) (9), offers nationally representative ($n > 20$ million) and contemporaneous data from computerised medical records (CMR) and biological specimen data – ideal for differentiated work in complex populations or remits (10). This paper documents our first steps to enhancing RCGP RSC surveillance of immunosuppressed patients by delivering an ORCHID-compatible and COVID-19 specific digital phenotype for 'Adult immunosuppression' – a phenotype being a set of rules that can be applied to CMR system data to identify or differentiate specific cohorts and/or episodes of interest (11). In the absence of a gold standard (5), we carried out a three-step process to deliver a testable pilot of this digital tool.

Leveraging international testing campaigns and their resultant research outputs (12), we first assessed whether COVID-19 mortality data was retrievable for the entire immunosuppressed

spectrum and, if so, whether visible clusters of risk existed to inform hierarchical stratification. We then utilised in-house clinical expertise to convert this conceptual hierarchy into a valueset compatible with CMR systems. Finally, in-house data expertise developed the logistical rules that could express this hierarchy within a data extract, allocating medical records into risk levels in accordance with the underpinning hierarchy while respecting key time and treatment dependencies and the need for mutual exclusivity. The present paper describes the outcomes of these exercises and the challenges and key learnings that emerged.

This work was conducted in the UK context, where *Immunisation against infectious disease* (Green Book) chapters set-out national vaccine policy and offer the textual phenotype for 'Adult Immunosuppression' used in this remit (13). However, unlike international analogues (14), Green Book does not supply detail on differential risk or stratify its characterisations of immunosuppression in any way. To the best of the authors' knowledge, this work is therefore the first to attempt to convert this specific resource into a format compatible with a hierarchical database search of the immunosuppressed and their COVID-19 outcomes. By doing so, it is our intention to create a feedback loop of real-world evidence to continually refine, augment and target Green Book COVID-19 guidelines for the benefit of immunosuppressed patients and the clinicians that treat them.

Method

Rapid scoping review

The rapid scoping review was performed as a hypothesis generating exercise. It was requested by stakeholders at the

United Kingdom Health Security Agency (UKHSA) to establish whether COVID-associated mortality could be extracted and compared across all immunosuppressed conditions and medications listed within *Immunisation against infectious disease* (Green Book) Chapter 14a (Box 1) (13) and if clusters of risk were visible at surface-level review. This request was made during the second year of the pandemic and leveraged the scale and specificity of the immunosuppressed COVID-19 literature base that had accumulated even by that early stage thanks to the proliferation of community testing (12). When attempted, there was insufficient data to complete this exercise for immunosuppressed Seasonal Influenza outcomes (Supplementary Materials).

This exercise applied the following PECO structure: Population – COVID-19 infection; Exposure – Adult Immunosuppression (as defined by Green Book Chapter 14a (13)); Comparator – Immunocompetent; Outcome – COVID-19 mortality. As summarised in Box 2, all medications, procedures and conditions listed in Green Book Chapter 14a were entered into Google Scholar as search terms along with mortality- and COVID-19-related language. Overlapping or duplicated terms were collapsed into common categories for simplicity and specific inclusion and exclusion criteria were applied. Searches were iterative and utilised the ‘relevant articles’ function of Google Scholar to redundancy. Search dates were January, 2020 to August, 2022 inclusive.

Exemplar studies on the COVID infection experience of each immunosuppressed category were selected - a selection process that considered sample size, research design, data completeness and efforts to minimise bias. Screening for exemplar studies occurred first by title and abstract before full text was assessed and data was extracted where relevant. This was also an iterative process, with exchanges to certain exemplar studies made when invalid or superior data was identified upon second review. An extraction form (populated further within) supported the collection of all necessary exemplar study characteristics, including Green Book category, Title, Author(s), Year, Country, Excess COVID Mortality (Effect Measure), 95% Confidence Interval, Proportion Mortality amongst All Case Types and Proportion Mortality in Hospitalised Cases. Quality assured by expert clinicians on an ongoing basis, observed clusters of effect sizes (denoting excess mortality between immunosuppressed categories and immunocompetent

counterparts) and mortality percentages informed our conceptual means of subdivision.

Developing the phenotype

As illustrated within Figure 1 (15), the RSC currently utilises a three-layered approach to preparing a novel phenotype for use within its data flows:

1. An ontological layer - a human-readable description of key concepts within, and relationships between, a given phenotype.
2. A coding layer – where conceptual phenotype layers are translated into their constituent clinical codes and thematic categories.
3. A logical data extract layer – where ontological rules, expressing the relationships between phenotype levels and their respective eligibility criteria, are converted into Structured Query Language (SQL) expressions to extract desired data from the overall pool(s) available.

This process was applied when converting the conceptual framework of immunosuppression that surfaced via rapid scoping review (our ontological layer) into a test-ready phenotype (test extract). The final phenotype schema will only be confirmed and published once the pilot phenotype has been tested and validated within the ORCHID. To be compatible with this server, the curation discussed here only utilised Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) diagnostic and Dictionary of Medicines and Devices (dm+d) prescription terms.

UKHSA currently commissions the University of Nottingham based Primary Care Information Services (PRIMIS) (16) to translate its risk groups into a SNOMED CT and dm+d reference set for its disease surveillance needs, including immunosuppression. PRIMIS versions 1.3.3 and 2.3 were consequently a key resource for our coding layer. However, PRIMIS is undifferentiated and is sensitive to redundancy as SNOMED CT and dm+d vocabularies are updated on a regular basis (monthly and weekly, respectively). Furthermore, PRIMIS clinical risk groups can be highly porous, with codes relevant for immunosuppression located in numerous other categories.

BOX 1 Immunosuppression in adult populations as defined by *Immunisation against infectious disease* (Green Book) Chapter 14a

Immunosuppression

Immunosuppression due to disease or treatment, including patients undergoing chemotherapy leading to immunosuppression, patients undergoing radical radiotherapy, solid organ transplant recipients, bone marrow or stem cell transplant recipients, HIV infection at all stages, multiple myeloma or genetic disorders affecting the immune system (e.g. IRAK-4, NEMO, complement disorder, SCID).

Individuals who are receiving immunosuppressive or immunomodulating biological therapy including, but not limited to, anti-TNF, alemtuzumab, ofatumumab, rituximab, patients receiving protein kinase inhibitors or PARP inhibitors, and individuals treated with steroid-sparing agents such as cyclophosphamide and mycophenolate mofetil.

Individuals treated with or likely to be treated with systemic steroids for more than a month at a dose equivalent to prednisolone at 20mg or more per day for adults.

Anyone with a history of haematological malignancy, including leukaemia, lymphoma, and myeloma.

Those who require long-term immunosuppressive treatment for conditions including, but not limited to, systemic lupus erythematosus, rheumatoid arthritis, inflammatory bowel disease, scleroderma, and psoriasis.

BOX 2 Example Search Terms and Inclusion & Exclusion Criteria for Rapid Literature Review

Search terms included in Google Scholar search:

COVID-19 AND “excess mortality” AND chemotherapy OR radiotherapy OR “solid organ transplant*” OR transplant OR “bone marrow” OR “stem cell” OR myeloma OR complement OR biologics OR “biological therap*” OR immunotherapy OR PID OR HIV OR AIDS OR steroid OR prednisolone OR malignancy OR lymphoma OR “rheumatoid arthri*” OR IBS OR scleroderma OR lupus OR psoriasis OR “steroid sparing” OR “anti-TNF” OR leukaemia OR “genetic disorder*” OR rituximab OR “protein kinase inhib*” OR “PARP*” OR cyclophosphamide OR “mycophenolate mofetil” OR “long term immuno*” OR “long-term immuno*” OR IRAK-4 OR NEMO OR SCID OR alemtuzumab OR ofatumumab.

Inclusion criteria:

- Study investigated the impact of COVID-19 on a specific immunosuppressed category (as defined by *Immunisation against infectious disease* [Green Book] Chapter 14a).
- Mortality data were provided in some form (proportion deceased and/or comparative mortality to immunocompetent via effect measure).
- Study reported adult data (over 18 years).

Exclusion criteria:

- Study did not investigate the impact of COVID-19 on a specific immunosuppressed category (as defined by *Immunisation against infectious disease* [Green Book] Chapter 14a).
- Mortality data were not provided in any form.
- Multiple/overlapping risk groups were assessed (e.g. pregnant/diabetic immunosuppressed patients).
- Immunosuppressives were investigated as a prospective treatment for COVID-19, not as a risk factor.
- Study included paediatric data (under 18 years).

As such, to create our valueset, all immunosuppressed PRIMIS terms had to be manually extracted, updated or scraped from wider risk groups using an in-house SNOMED CT Concept ID tool (described elsewhere (17)). Supertype and theme variables specific to the ORCHID TRE, known as Themes Access Dynamic Data Services [TADDS] variables, were also included in our coding layer to improve the comprehensiveness of our valueset and to reduce its vulnerability to redundancy [ibid]. Each with their own ID, TADDS variables denote specific disease or medication types and can be easily updated using Expression Constraint Language (ECL). Once completed, this valueset could then be subdivided in accordance with rapid scoping review findings; this final immunosuppressed

valueset (provided in [Supplementary Materials](#)) was then cross-checked for accuracy and comprehensiveness against Chapter 14a specifications. Multiple expert clinicians (DK, AF, RW, GJ) quality assured this process.

The logical extract model aimed to incorporate the time conditions specified in Chapter 14a immunosuppressed criteria (13) and relevant literature (18) into our pilot phenotype. Furthermore, hierarchical principles were applied to ensure that, once tested within real-world data, our pilot phenotype would eliminate multi-level allocation of medical records. This mutual exclusivity ensured patients would only be allocated into, and contribute their data to, the highest risk level they were eligible for.

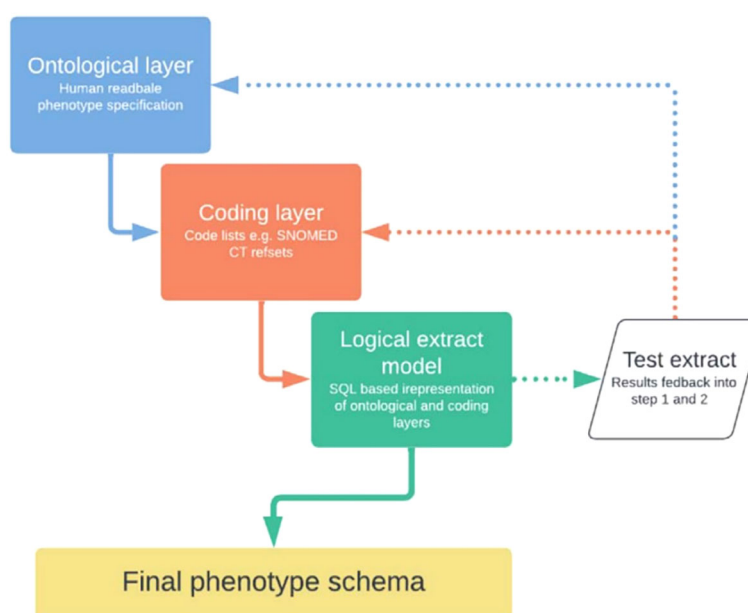


FIGURE 1
Three-layered approach to instituting a novel phenotype in RSC data.

Results

Rapid scoping review

Due to the hypothesis-generating nature of this work, its iterative nature and the time-sensitivity involved, data was not retained on the number of articles returned, reviewed or excluded.

Once overlapping or synonymous immunosuppressed categories were collapsed, twelve categories were left to be populated with their respective exemplar cohort studies (19–27). The data extraction form that supported this process is provided in [Table 1](#) below. Exemplar studies that met search criteria, reported excess mortality between immunosuppressed subgroups and their immunocompetent comparators are reported in [Figure 2](#). Here, ‘Effect Size’ reflects the variety of measures that were extracted, including Odds, Hazard and Risk Ratios. The mathematical equivalence of these measures, as argued by Symons & Moore (2002) (28), justified our decision to report unadjusted data via this umbrella term.

Exemplar study estimates of COVID-19 mortality risk varied considerably across the immunosuppressed population. Higher versus lower risk categories are identifiable, with agreement between mortality percentages and effect size estimates on the elevated vulnerability of haematological malignancy patients and recipients of systemic steroids and of minimal excess mortality seen amongst HIV, chemotherapy and radiotherapy patients. Disparities in risk can also be seen within immunosuppressed condition and medication types (solid tumours versus haematological malignancies and biologics versus cancer therapies, respectively). There were notable discrepancies between the risk profiling that emerged from mortality percentages versus effect size estimates, however. The elevated effect size of solid organ transplant recipients is not borne out in percentage data, for example.

Due to the heterogeneity of exemplar studies’ designs and sample sizes, effect measure-based estimates of COVID-19 mortality were prioritised over percentage data. This comparative metric standardised between-category estimates and made for fairer comparisons than would be achieved using absolute values.

The following three patient clusters proved most valuable for dividing immunosuppressed COVID-19 data in the first instance: 1) those with bone marrow compromising conditions, 2) those receiving ongoing immunosuppressing treatments and/or procedures and 3) those with primary or acquired immunodeficiencies. To prevent oversimplification, RSC clinicians (DK, AF) expanded upon this initial tripartite model to create the five-part phenotype elaborated upon in the next section of reporting. Here, dedicated strata were created for solid organ transplant recipients and immunosuppressive regimens related and unrelated to cancer in recognition of their unique risk profile and easy confusion in large databases respectively (the same immunosuppressive agents can be prescribed as cancer and non-cancer therapies). [Table 2](#) provides brief descriptions of these five levels and their respective immunosuppressive conditions or treatment regimens.

Developing the phenotype

The test valueset is provided in [Supplementary Materials](#). As illustrated within [Figure 3](#), a combination of PRIMIS terms and cancer-specific TADDS IDs translated our conceptual five-part phenotype into this SNOMED and dm+d compatible format. Due to a major SNOMED CT update over the course of the curation process, new terms for newly redundant or retired codes had to be identified and manually inserted. As part of this process, PRIMIS sources also went through additional cross-checks for immunosuppression-eligible codes located in broader risk categories; several transplantation codes were scraped from chronic kidney and chronic liver disease categories, for example. Finally, the cancer diagnosis TADDS variables that differentiated between Non-Cancer Immunosuppressive Treatments and Recent Cancer Treatment layers were curated and quality assured by colleagues within the Nuffield Department of Primary Care Health Sciences, including RSC clinical and data team members (DK, GJ).

Operationalised with SQL, ontological rules for hierarchical mutual exclusivity, time sensitivity and diagnosis-treatment pairing stratify this valueset into our desired pilot phenotype for ‘Adult Immunosuppression’ for onward use in COVID-19 surveillance. Rules currently follow Green Book time directives (13) and best practice from contemporary literature (18) to exclude those that would be considered insufficiently immunosuppressed during extraction (e.g. legacy cancer patients and those whose immunosuppressive regimen discontinued greater than 6 months prior to infection date). This pilot phenotype is now being tested within ORCHID data; results will be reported imminently. Any difficulties, inaccuracies or misclassifications that surface during this test will be used to refine the phenotype prior to its publication as our final phenotype schema within BioPortal (29), PhenoFlow (30) or an equivalent.

Discussion

This paper outlines the development of a risk-stratified, test-ready phenotype for ‘Adult Immunosuppression’ for COVID-19 surveillance dataflows. To the best of our knowledge, this work is the only to operationalise UK specifications for immunosuppression in a hierarchical manner. Utilising COVID-19 mortality as our means of subdivision, this pilot phenotype represents a novel, clinically relevant and medical health record compatible basis for identifying the immunosuppressed in CMR repositories and appraising their COVID-19 outcomes by vulnerability level. Once implemented, this has significant prospective value for COVID-19 surveillance and, if validated, patient triage and treatment prioritisation. Due to the idiosyncrasies of COVID-19 (31), transferability of this phenotype to other disease domains can only be speculated at this stage, however.

TABLE 1 Completed data extraction form of exemplar rapid scoping review cohort studies comparing COVID-19 associated mortality between immunosuppressed categories and immunocompetent comparators.

Green Book Category	Reference	Year	Country	Immunosuppressed cases (n)	Immunosuppressed mortality (n)	Comparator cases (n)	Comparator mortality (n)	Effect Size	LCI	UCI	All case mortality* (%)	Hospitalised mortality** (%)
Chemotherapy	Lee et al. (19)	2020	UK	281	75	272	92	1.18	(0.81)	1.72)	Not reported	27.0%
Radical radiotherapy	Lee et al. (19)	2020	UK	76	18	272	92	0.65	(0.36)	1.18)	Not reported	24.0%
Solid Organ Transplantation	Sun et al. (20)	2021	USA	11392	875	1426984	23830	3.38	(3.35)	3.41)	7.68%	Not reported
Bone Marrow/Stem Cell Transplantation	Ljungman et al. (21)	2021	Global	382	95	Not reported	Not reported	Not reported	Not reported	Not reported	25.20%	Not reported
HIV	Hadi et al	2021	USA	404	20	404	15	1.33	(0.69)	2.57)	4.95%	Not reported
Multiple Myeloma	Chari et al. (23)	2020	Global	617	203	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported	33.0%
Primary Immunodeficiencies	Shields et al. (24)	2021	UK	310	55	Not reported	Not reported	Not reported	Not reported	Not reported	17.70%	Not reported
Biologics	Suárez-García et al. (25)	2021	Spain	183	49	11095	2143	1.97	(1.33)	2.91)	Not reported	26.7%
Steroid sparing agents	Yousaf et al	2020	Global	213	13	213	15	0.91	(0.68)	1.22)	6.10%	Not reported
Systemic steroids	Suárez-García et al. (25)	2021	Spain	570	202	11095	2143	2.16	(1.8)	2.61)	Not reported	35.4%
Haematological malignancies	Passamonti et al. (27)	2020	Italy	536	198	239627	33498	2.04	(1.77)	2.34)	37.00%	Not reported
Long Term Immunosuppressives	Suárez-García et al. (25)	2021	Spain	394	109	11095	2143	2.06	(1.64)	2.6)	Not reported	27.7%

* Percentage of deaths recorded amongst all COVID cases.

** Percentage of deaths recorded exclusively amongst cases hospitalised with COVID.

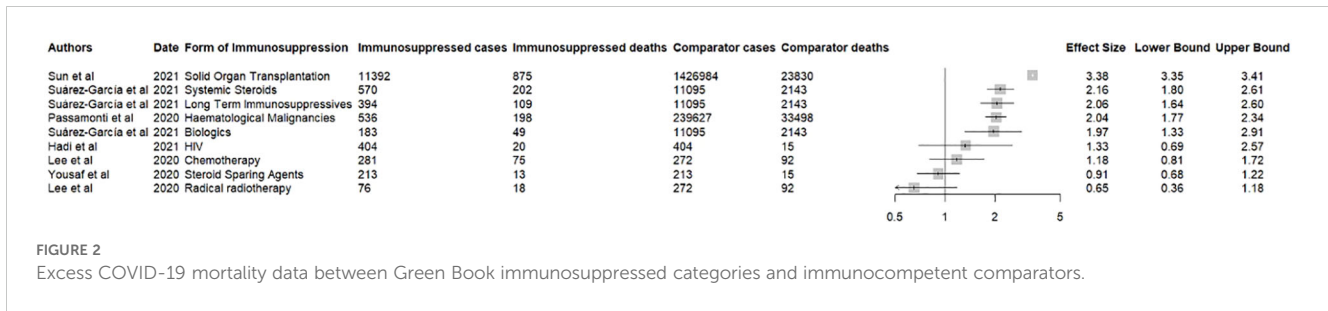


FIGURE 2 Excess COVID-19 mortality data between Green Book immunosuppressed categories and immunocompetent comparators.

Our pilot phenotype has already benefitted from the oversight and quality assurance of the in-house clinical team of the RSC. Work is now underway to assess its capabilities in real-world data; feasibility exercises within ORCHID will be published in due course. As per Methods, findings will be used to either validate or feedback and refine the pilot phenotype.

Despite the prospective value of this pilot phenotype, it must be restated that this work was intended as a hypothesis generating exercise; its explorative and rapid nature makes it inappropriate for general use and the valueset provided requires further validation before we can recommend this.

In the first instance, risk stratification was based on single-study estimates for COVID-19 mortality and the practical requirements of CMR extraction. The superficiality of the rapid scoping exercise that procured this data warrants criticism – its transparency and replicability were both undermined by failing to pre-publish a protocol and retain precise screening data. Furthermore, the heterogeneity of exemplar studies’ designs and sample sizes confounded comparisons between absolute mortality outcomes. While effect measure data (denoting excess mortality as compared to immunocompetent counterparts) could compensate by providing a more standardised form of comparison, they were not reported for all subgroups. A systematic replication and meta-analysis were recently conducted (32) to address all these issues and to test the rigour of pilot phenotype levels. Despite the general concordance between the two exercises, the results of this

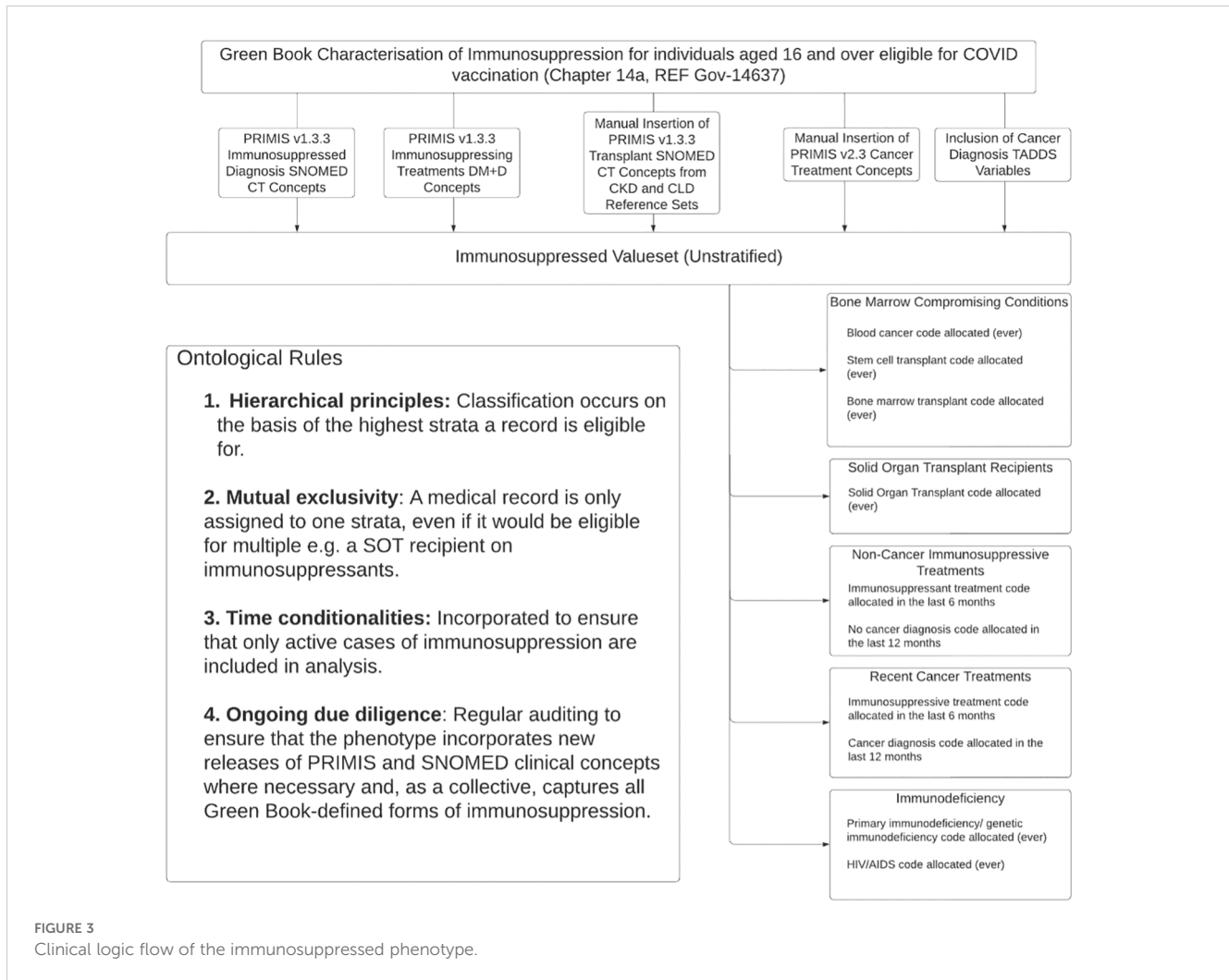
systematic work suggest some inaccuracies in our pilot phenotype – most notably the collapsing of acquired with primary immunodeficiencies given the generally elevated risk profile of the latter (33) – and, thanks to its global scope, some key geographical distinctions. COVID infection carries very different risks for the HIV patient with access to antiretroviral therapy versus without, for example (34). An eDelphi study is currently underway to leverage this evidence to produce a more definitive COVID-19 risk stratified phenotype for ‘Adult Immunosuppression’ (35).

The expression of our pilot phenotype is also not without its limitations. In the first instance, the mutual exclusivity that was codified between phenotype levels does not exist in reality. Immunosuppressive conditions rarely go untreated, for example. This will likely make it difficult to demarcate mortality risks between diagnoses and their medications. Furthermore, the extensional nature of the pilot phenotype, manually curating levels on a code-by-code basis, makes it susceptible to inaccuracies and redundancies as both PRIMIS and its underpinning SNOMED and dm+d clinical concepts update on a regular basis. While rigorous quality assurance efforts and intentional use of supertype terms and TADDS variables makes us confident that our valueset is both reflective of the diversity of Green Book immunosuppressed terms and somewhat resilient to redundancy, future iterations of this work would benefit from intensional – rule-based, machine-readable and more readily-sharable – curation methods. In the spirit of open science, this will be looked to as the standard for all future RSC phenotypes (17).

Finally, the implementation of the pilot phenotype may prove challenging in certain areas. Automating patient allocation via time and dual treatment-diagnosis is complex and may lead to inaccuracies when applied to computerised medical records; conditional levels, most notably those that are medication-specific, are far harder to extract than their unconditional and diagnostic-specific counterparts. For example, 12-month caps on cancer diagnostic codes and 6-month minimums for immunosuppressive medications were incorporated to prevent the inclusion of legacy cancers or low-level immunosuppression respectively. However, such crude constraints may lead to oversimplification, under sampling or misclassification in practice – especially when considering well-documented issues in keeping primary care records and prescription data updated (36). These caps are also at odds with more recent clinical logic to identify immunosuppressed patients (37); this has extended cancer

TABLE 2 Composition of Pilot Phenotype for ‘Adult Immunosuppression’, stratified by COVID-19 mortality risk.

Phenotype Level	Composition
1. Bone Marrow Compromising Conditions	Haematological malignancy and Bone Marrow/Stem Cell Transplantation
2. Solid Organ Transplant Recipients	Solid Organ Transplantation
3. Non-Cancer Immunosuppressive Treatments	Immunosuppressives unrelated to cancer
4. Recent Cancer Treatments	Systemic cancer treatments including cancer-related immunotherapy, chemotherapy and radical radiotherapy
5. Immunodeficiency	Genetic immunodeficiencies and HIV/AIDS



treatment caps to 12 months and 24 months for chemo/radiotherapy and Chimeric Antigen Receptor T-cells (CAR-T) therapy respectively. This work also failed to operationalise Green Book dose dependencies for immunosuppression (e.g. equivalence to ≥ 20 mg of prednisolone) which is a major limitation. In acknowledgment of the challenges of automating drug-inclusive phenotypes for immunosuppression, future iterations of this work may wish to consider condition-specific phenotyping or binary alternatives (higher vs lower risk) – either precluding medication codes altogether or simplifying allocation. Clinical consensus will be built on these topics via the eDelphi exercise previously mentioned (35).

Conclusion

Despite its hypothesis-generating intentions, this pilot phenotype for ‘Adult Immunosuppression’ from a COVID-19 lens is the first we are aware of to automate the identification and

risk-stratification of this ill-defined and highly vulnerable population. Its development and curation have already returned valuable insights for those working to enhance COVID-19 surveillance and its first trial in real-world data is underway. This validation is urgently needed, though concordances between the rapid scoping review described here and its systematic replication (32) are encouraging. Results of both this and the eDelphi study (35) to build international consensus on immunosuppressed definitions and COVID risks will be utilised to create and publish a final schema phenotype of ‘Adult immunosuppression’ for COVID-19 dataflows. Until this point, the materials provided in this paper should not be considered ready for general use.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary Material](#). Further inquiries can be directed to the corresponding author.

Author contributions

ML: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Visualization, Writing – original draft, Writing – review & editing. DK: Data curation, Writing – review & editing. GJa: Data curation, Writing – review & editing. AF: Data curation, Writing – review & editing. RW: Data curation, Writing – review & editing. GJi: Data curation, Writing – review & editing. JO: Methodology, Supervision, Writing – review & editing. DS: Writing – review & editing. HW: Writing – review & editing. MJ: Supervision, Writing – review & editing. LL: Supervision, Writing – review & editing. FH: Supervision, Writing – review & editing. Sd: Supervision, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work is funded by the Medical Research Council iCASE Studentship: financial contributions are made by the MRC and EMIS Health.

References

- Memoli MJ, Athota R, Reed S, Czajkowski L, Bristol T, Proudfoot K, et al. The natural history of influenza infection in the severely immunocompromised vs nonimmunocompromised hosts. *Clin Infect Dis*. (2013) 58(2):214–24. doi: 10.1093/cid/cit725
- Belsky JA, Tullius BP, Lamb MG, Sayegh R, Stanek JR, Auletta JJ. Covid-19 in immunosuppressed patients: A systematic review of cancer, hematopoietic cell and solid organ transplant patients. *J Infection*. (2021) 82(3):329–38. doi: 10.1016/j.jinf.2021.01.022
- Tassone D, Thompson A, Connell W, Lee T, Ungaro R, An P, et al. Immunosuppression as a risk factor for COVID-19: A meta-analysis. *Internal Med J*. (2021) 51(2):199–205. doi: 10.1111/imj.15142
- Cortegiani A, Madotto F, Gregoretti C, Bellani G, Laffey JG, Pham T, et al. Immunocompromised patients with acute respiratory distress syndrome: Secondary analysis of the Lung Safe Database. *Crit Care*. (2018) 22(1):157. doi: 10.1186/s13054-018-2079-9
- Greenberg JA, Hohmann SF, Hall JB, Kress JP, David MZ. Validation of a method to identify immunocompromised patients with severe sepsis in administrative databases. *Ann Am Thorac Soc*. (2016) 13(2):253–8. doi: 10.1513/annalsats.201507-415bc
- Shoham S, Batista C, Ben Amor Y, Ergonul O, Hassanain M, Hotez P, et al. Vaccines and therapeutics for immunocompromised patients with covid-19. *eClinicalMedicine*. (2023) 59:101965. doi: 10.1016/j.eclinm.2023.101965
- Department of Health and Social Care. *Higher-Risk Patients Eligible for COVID-19 Treatments: Independent Advisory Group Report* (2022). GOV.UK. Available online at: <https://www.gov.uk/government/publications/higher-risk-patients-eligible-for-covid-19-treatments-independent-advisory-group-report> (Accessed 1 August 2024).
- Whitaker HJ, Tsang RS, Byford R, Andrews NJ, Sherlock J, Pillai PS, et al. Pfizer-BioNTech and Oxford AstraZeneca COVID-19 vaccine effectiveness and immune response among individuals in clinical risk groups. *J Infection*. (2022) 84:675–83. doi: 10.1016/j.jinf.2021.12.044
- de Lusignan S, Jones N, Dorward J, Byford R, Liyanage H, Briggs J, et al. The oxford royal college of general practitioners clinical informatics digital hub: protocol to develop extended COVID-19 surveillance and trial platforms. *JMIR Public Health Surveill*. (2020) 6:e19773. doi: 10.2196/19773
- Leston M, Elson WH, Watson C, Lakhani A, Aspden C, Bankhead CR, et al. Representativeness, vaccination uptake, and COVID-19 clinical outcomes 2020–2021 in the UK Oxford-royal college of general practitioners research and surveillance network: cohort profile summary. *JMIR Public Health Surveill*. (2022) 8:e39141. doi: 10.2196/39141
- Brandt PS, Kho A, Luo Y, Pacheco JA, Walunas TL, Hakonarson H, et al. Characterizing variability of electronic health record-driven phenotype definitions. *J Am Med Inform Assoc*. (2023) 30:427–37. doi: 10.1093/jamia/ocac235

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2024.1491565/full#supplementary-material>

- Furstenau LB, Rabaioli B, Sott MK, Cossul D, Bender MS, Farina EMJDM, et al. A bibliometric network analysis of coronavirus during the first eight months of COVID-19 in 2020. *Int J Environ Res Public Health*. (2021) 18:952. doi: 10.3390/ijerph18030952
- UKHSA. (2023) Covid-19: The Green Book, Chapter 14A, GOV.UK. Available online at: <https://www.gov.uk/government/publications/covid-19-the-green-book-chapter-14a?UID=67444390920231028203857> (Accessed 06 November 2023).
- Australian Government Department of Health and Aged Care. Vaccination for people who are immunocompromised. In: *The Australian Immunisation Handbook* (2022). Available at: immunisationhandbook.health.gov.au/contents/vaccination-for-special-risk-groups/vaccination-for-people-who-are-immunocompromised. (Accessed August 14, 2024).
- de Lusignan S. In this issue: Ontologies a key concept in informatics and key for open definitions of cases, exposures, and outcome measures. *J Innov Health Inform*. (2015) 22:170. doi: 10.14236/jhi.v22i2.170
- PRMIS, University of Nottingham. COVID-19 vaccination uptake reporting specification(2021). Available online at: <https://www.nottingham.ac.uk/prmis/covid-19/covid-19.aspx>. (Accessed August 14, 2024)
- Jamie G, Elson WH, de Lusignan S, Kar D, Wimalaratna R, Hoang U, et al. Phenotype Execution and Modelling Architecture (PhEMA) to support disease surveillance and real-world evidence studies: English sentinel network evaluation. *medRxiv (Cold Spring Harbor Laboratory)*. (2023) 7(2):o0ae034. doi: 10.1101/2023.11.21.23298758
- Lee LYW, Starkey T, Ionescu MC, Little M, Tilby M, Tripathy AR, et al. Vaccine effectiveness against COVID-19 breakthrough infections in patients with cancer (UKCCEP): a population-based test-negative case-control study. *Lancet Oncol*. (2022) 23:748–57. doi: 10.1016/s1470-2045(22)00202-9
- Lee LY, Cazier JB, Angelis V, Arnold R, Bisht V, Campton NA, et al. Covid-19 mortality in patients with cancer on chemotherapy or other anticancer treatments: A prospective cohort study. *Lancet*. (2020) 395:1919–26. doi: 10.1016/s0140-6736(20)31173-9
- Sun J, Patel RC, Zheng Q, Madhira V, Olex AL, Islam JY, et al. COVID-19 disease severity among people with HIV infection or solid organ transplant in the United States: A nationally-representative, multicenter, observational cohort study. *medRxiv*. (2021) 28. doi: 10.1101/2021.07.26.21261028
- Ljungman P, de la Camara R, Mikulska M, Tridello G, Aguado B, Zahrani MA, et al. COVID-19 and stem cell transplantation; results from an EBMT and GETH multicenter prospective survey. *Leukemia*. (2021) 35:2885–94. doi: 10.1038/s41375-021-01302-5
- Hadi YB, Naqvi SFZ, Kupec JT, Sarwari AR. Characteristics and outcomes of COVID-19 in patients with HIV: a multicenter research network study. *AIDS*. (2020) 34:F3–8. doi: 10.1097/QAD.0000000000002666

23. Chari A, Samur MK, Martinez-Lopez J, Cook G, Biran N, Yong K, et al. Clinical features associated with COVID-19 outcome in multiple myeloma: first results from the International Myeloma Society data set. *Blood*. (2020) 136:3033–40. doi: 10.1182/blood.202008150
24. Shields AM, Anantharachagan A, Arumugakani G, Baker K, Bahal S, Baxendale H, et al. Outcomes following SARS-CoV-2 infection in patients with primary and secondary immunodeficiency in the United Kingdom. *Res Square*. (2021). 209(3):247–58 doi: 10.21203/rs.3.rs-920110/v1
25. Suárez-García I, Perales-Fraile I, González-García A, Muñoz-Blanco A, Manzano L, Fabregate M, et al. In-hospital mortality among immunosuppressed patients with COVID-19: Analysis from a national cohort in Spain. *PLoS One*. (2021) 16:e0255524. doi: 10.1371/journal.pone.0255524
26. Yousaf A, Gayam S, Feldman S, Zinn Z, Kolodney M. Clinical outcomes of COVID-19 in patients taking tumor necrosis factor inhibitors or methotrexate: A multicenter research network study. *J Am Acad Dermatol*. (2021) 84:70–5. doi: 10.1016/j.jaad.2020.09.009
27. Passamonti F, Cattaneo C, Arcaini L, Bruna R, Cavo M, Merli F, et al. Clinical characteristics and risk factors associated with COVID-19 severity in patients with haematological Malignancies in Italy: a retrospective, multicentre, cohort study. *Lancet Haematol*. (2020) 7:e737–45. doi: 10.1016/S2352-3026(20)30251-9
28. Symons MJ, Moore DT. Hazard rate ratio and prospective epidemiological studies. *J Clin Epidemiol*. (2002) 55:893–9. doi: 10.1016/s0895-4356(02)00443-2
29. Goldman JD, Robinson PC, Uldrick TS, Ljungman P. COVID-19 in immunocompromised populations: implications for prognosis and repurposing of immunotherapies. *J ImmunoTherapy Cancer*. (2021) 9:e002630. doi: 10.1136/jitc-2021-002630
30. Common Workflow Language. *Phenoflow | Common Workflow Language Viewer* (2020). Available online at: kclhi.org/phenoflow/ (Accessed 28 Oct. 2024). *Kclhi.org*.
31. National Center for Biomedical Ontology. *Welcome to the NCBO BioPortal | NCBO BioPortal* (2024). Available online at: bioportal.bioontology.org (Accessed August 14, 2024)
32. Leston M, Elson W, Ordóñez-Mena JM, Kar D, Whitaker H, Joy M, et al. Disparities in COVID-19 mortality amongst the immunosuppressed: A systematic review and meta-analysis for enhanced disease surveillance. *J Infection*. (2024) 88:106–10. doi: 10.1016/j.jinf.2024.01.009
33. SeyedAlinaghi S, Karimi A, Barzegary A, Mojdeganlou H, Vahedi F, Mirghaderi SP, et al. COVID-19 mortality in patients with immunodeficiency and its predictors: a systematic review. *Eur J Med Res*. (2022) 27(1):195. doi: 10.1186/s40001-022-00824-7
34. Kassarjee R, Davies MA, Ngwenya O, Osei-Yeboah R, Jacobs T, Morden E, et al. COVID-19 among adults living with HIV: correlates of mortality among public sector healthcare users in Western Cape, South Africa. *J Int AIDS Soc*. (2023) 26:e26104. doi: 10.1002/jia2.26104
35. Leston M, Lee L, Ordóñez-Mena José, Joy M, de Lusignan S, Hobbs R. Defining and risk-Stratifying immunosuppression (the DESTINIES study): protocol for an electronic Delphi study. *JMIR Res Protoc*. (2024) 13:e56271. doi: 10.2196/56271
36. McCutchan G, Ordóñez-Mena J, Joy M, de Lusignan S, Hobbs R, McInnes I, et al. Assessing, updating and utilising primary care smoking records for lung cancer screening. *BMC Pulm Med*. (2023) 23:445. doi: 10.1186/s12890-023-02746-4
37. Services. *Rule logic - NHS England Digital* (2023). NHS England Digital. Available online at: <https://digital.nhs.uk/services/cohorting-as-a-service-caas/covid-at-risk-cohort/rule-logic> (Accessed 13 Aug. 2024).