**Keywords for filtering notes with social determinants of health (SDoH)**

'discharge summary', 'social worker', 'social history', 'income', 'employment', 'unemployment', 'financial', 'education', 'house', 'housing', 'crime', 'violence', 'physical activity', 'neighborhood', 'neighbourhood', 'community', 'poverty', 'occupation', 'indigence', 'homeless', 'shelter’, 'socioeconomic', 'lives with', 'living with', 'resideswith', 'worked in', 'work as', 'works as', 'working in', 'working at', 'working as', 'employed as', 'employed in', 'retired', 'accompanied by', 'here with', 'disability', 'disabled', 'food insecurity', 'transportation need', 'physical abuse', 'physical activity', 'sexual abuse', 'well developed'

**Table S1. Summary of attributes for social determinants of health concepts.**

|  |  |
| --- | --- |
| **SDoH concept** | **Attributes** |
| Tobacco\_use | Substance\_use\_status: e.g., current smokerSmoking\_type: type of tobacco use, e,g., cigarettesSmoking\_freq\_ppd: pack per daySmoking\_freq\_sy: years of smokingSmoking\_freq\_py: pack yearSmoking\_freq\_qy: years of quit smokingSmoking\_other: other information related with tobacco use |
| Alcohol use | Substance\_use\_status: e.g., “yes”, “no”Alcohol\_freq: e.g., “25.0 standard drinks of alcohol per week”Alcohol\_type: e.g., “wine” Alcohol\_other: other information related with alcohol use |
| Drug use | Substance\_use\_status: e.g., “yes”, “no”, “severe”Drug\_freq: frequency of drug useDrug\_type: e.g., “cocaine”Drug\_other: other inforamtion related with drug use |
| Sexual activity | Sdoh\_status: e.g., “yes”, “never”, “no”Partner: “female”, “male” Protection: protections of sexual activity  |
| Marital status | N/A |
| Education | N/A |
| Language | N/A |
| Employment | Occupation: job position, title, specialty, e.g., “in the meat department”Employment\_status: employment status such as “employed”, “unemployed”, “retired”Employment\_location: location of employment, e.g., “works in a lab” |
| Gender | N/A |
| Race | N/A |
| Ethnicity | N/A |
| Physical activity | Physical\_act: physical activity such as “running”, “exercise” Sdoh\_freq: e.g., “7 days per week”Sdoh\_status: e.g., “yes”, “no”, “currently not exercising” |
| Living condition | Living\_Condition: e.g., “homeless”, “shelter”, “live with mom” |
| Living supply | Living\_supply: food and water resource, and other supplies required for livingSdoh\_status: “yes”, “no”, “never”, indicating whether patients have supply issues |
| Financial constraint  | Financial\_constrain: descriptions for financial constrain, e.g., “difficulty paying for basics?”Sdoh\_status: “yes”, “no”, “not hard at all” |
| Abuse | Abuse: general/physical/mental abuse |
| Transportation | Transportation: “lack of transportation for medical needs?” Sdoh\_status: e.g., “yes”, “no” |
| Social cohesion | Social\_cohesion: Any social activity including person-person or person-organization interactions, e.g., “talks on phone”, “attend religious service”Sdoh\_status: “yes”, “no”, “never”, “sometimes”Sdoh\_freq: e.g., “twice a week”, “2 to 4 times a year” |
| SDoH ICD | SDoH\_ICD: ICD-10 codes appeared in notes associated with SDH (e.g., ICD-10 Z55-Z65) |

**Table S2. Detailed results for GatorTron to identify social determinants of health concepts and attributes.**

|  |  |  |
| --- | --- | --- |
|  | **Strict** | **Lenient** |
|  | **Prec.** | **Rec.** | **F1** | **Prec.** | **Rec.** | **F1** |
| Race | 0.9091 | 1 | 0.9524 | 0.9091 | 1 | 0.9524 |
| Abuse | 0.9412 | 0.9091 | 0.9249 | 0.9412 | 0.9091 | 0.9249 |
| Gender | 0.933 | 0.9435 | 0.9382 | 0.933 | 0.9435 | 0.9382 |
| Partner | 0.8824 | 0.8824 | 0.8824 | 1 | 1 | 1 |
| Sex\_act | 0.9792 | 0.9895 | 0.9843 | 0.9792 | 0.9895 | 0.9843 |
| Drug\_use | 0.9808 | 0.9189 | 0.9488 | 0.9808 | 0.9189 | 0.9488 |
| Language | 1 | 1 | 1 | 1 | 1 | 1 |
| Drug\_type | 0.8182 | 0.75 | 0.7826 | 0.9091 | 0.8333 | 0.8696 |
| Sdoh\_freq | 0.9773 | 1 | 0.9885 | 0.9773 | 1 | 0.9885 |
| Drug\_freq | 1 | 1 | 1 | 1 | 1 | 1 |
| Education | 0.9268 | 0.9344 | 0.9306 | 0.9593 | 0.9672 | 0.9633 |
| Ethnicity | 0.875 | 0.875 | 0.875 | 1 | 1 | 1 |
| Occupation | 0.7604 | 0.7849 | 0.7725 | 0.8617 | 0.871 | 0.8663 |
| Drug\_other | 0 | 0 | 0 | 1 | 0.3333 | 0.5 |
| Protection | 0.6667 | 0.5 | 0.5714 | 0.6667 | 0.5 | 0.5714 |
| Alcohol\_use | 0.9217 | 0.9138 | 0.9177 | 0.9217 | 0.9138 | 0.9177 |
| Tobacco\_use | 0.7 | 0.75 | 0.7241 | 0.7 | 0.75 | 0.7241 |
| Sdoh\_status | 0.9573 | 0.9515 | 0.9544 | 0.9634 | 0.9576 | 0.9605 |
| Physical\_act | 0.9787 | 0.9583 | 0.9684 | 0.9787 | 0.9583 | 0.9684 |
| Alcohol\_freq | 0.8103 | 0.7705 | 0.7899 | 1 | 0.918 | 0.9573 |
| Smoking\_type | 0.9931 | 0.9863 | 0.9897 | 0.9931 | 0.9863 | 0.9897 |
| Alcohol\_other | 0.55 | 0.6471 | 0.5946 | 0.65 | 0.7647 | 0.7027 |
| Living\_supply | 0.9652 | 0.9737 | 0.9694 | 0.9652 | 0.9737 | 0.9694 |
| Gender\_status | 0.4 | 0.8 | 0.5333 | 0.625 | 1 | 0.7692 |
| Marital\_status | 0.9633 | 0.9292 | 0.9459 | 1 | 0.9646 | 0.982 |
| Transportation | 1 | 0.9744 | 0.987 | 1 | 0.9744 | 0.987 |
| Social\_cohesion | 0.9889 | 0.978 | 0.9834 | 1 | 0.989 | 0.9945 |
| Smoking\_freq\_qy | 0.6923 | 0.7714 | 0.7297 | 0.8462 | 0.9429 | 0.8919 |
| Smoking\_freq\_py | 0.8667 | 0.9286 | 0.8966 | 0.8667 | 0.9286 | 0.8966 |
| Smoking\_freq\_sy | 0.9804 | 0.9615 | 0.9709 | 0.9804 | 0.9615 | 0.9709 |
| Smoking\_freq\_ppd | 0.9714 | 0.9855 | 0.9784 | 0.9857 | 1 | 0.9928 |
| Living\_condition | 0.5714 | 0.5714 | 0.5714 | 0.7857 | 0.7857 | 0.7857 |
| Employment\_status | 0.2 | 0.0833 | 0.1176 | 0.6 | 0.25 | 0.3529 |
| Smoking\_freq\_other | 0.5 | 0.5714 | 0.5333 | 0.625 | 0.7143 | 0.6667 |
| Financial\_constrain | 0.8095 | 0.8947 | 0.85 | 0.8095 | 0.8947 | 0.85 |
| **Overall (micro)** | **0.9197** | **0.9146** | **0.9172** | **0.9434** | **0.9359** | **0.9396** |



**Figure S1.** An overview of the 7 social determinants of health classes and 38 subclasses.