

**Appendix.** Rule Set Used for Algorithm 1

Entire words <sup>a</sup>	Wildcard words <sup>b</sup>	Fuzzy words <sup>c</sup>	Words in specific fields <sup>d</sup>	Proximity searches <sup>e</sup>
Viral	*itis	Hemorrhagic	<i>Undetermined</i> in cause of death field	<i>immune</i> within 2 words on either side of <i>syndrome</i>
AIDS	septic*	Thrombocytopenia	<i>Unknown</i> in cause of death field	
Sepsis	necro*	Hemolytic		
	*pneumo*	Tuberculosis		
	gangren*			
	bacter*			
	infect*			
	mening*			
	absces*			

<sup>a</sup>Entire words are searches for the word as written.

<sup>b</sup>Wildcard words are searches for the string of letters as written, with the asterisk representing any letter or letters. For example, *gangren\** would represent gangrene and gangrenous.

<sup>c</sup>Fuzzy words are searches for the word as written, but one letter may be misspelled, missing, inserted, or transposed, and the search will still identify that word. For example, a fuzzy search for *hemorrhagic* would identify hemorrhagic, haemorrhagic, and hemorhagic.

<sup>d</sup>Words in specific fields are searches for the words as written in specific database fields, in this case the cause of death fields.

<sup>e</sup>Proximity searches are searches for words in a given proximity to other words. For example, *immune* within 2 words on either side of *syndrome* would identify immune syndrome, immune deficiency syndrome, and syndrome of immune deficiency.