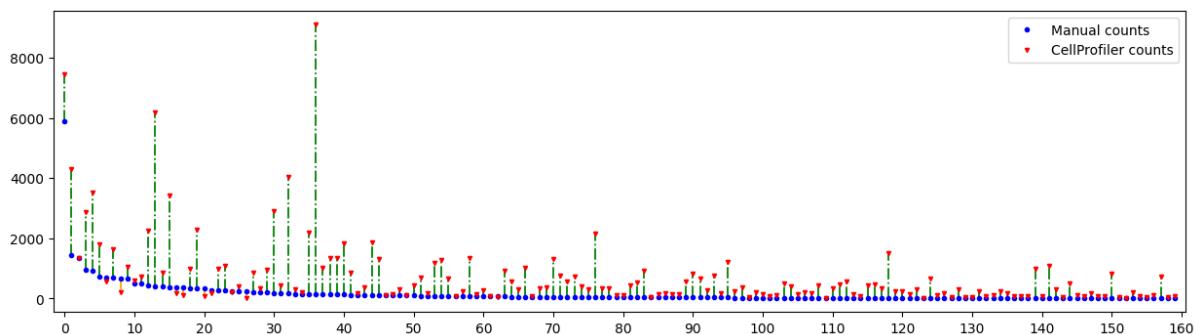


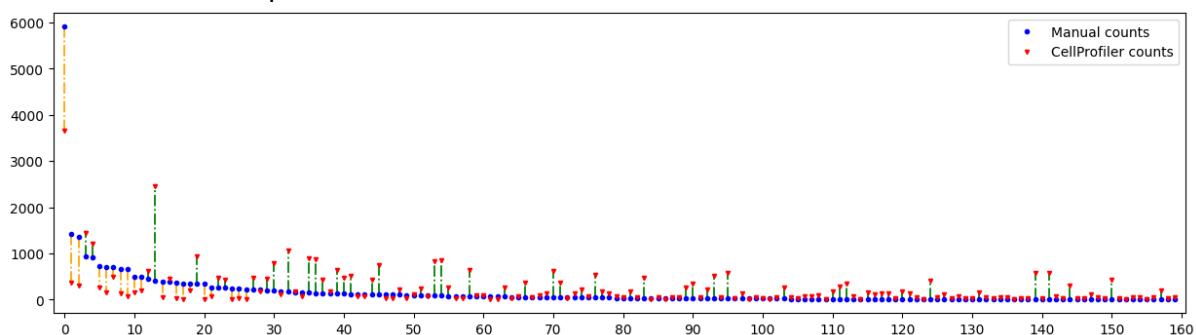
The CellProfiler pipeline for CD66b has been improved iteratively. For the CD66b staining we have 160 whole scale images (WSIs) where the CD66b-positive cells have been counted manually. The performance of each pipeline is determined by comparing the counts from CellProfiler to the manual counts.

Below you can see an overview of each iteration of the CellProfiler pipeline compared to the manual counts. In each figure the blue dots are the manual counted cells and the red dots the CellProfiler counts. For each iteration, the correlation between the CellProfiler and manual counts is given

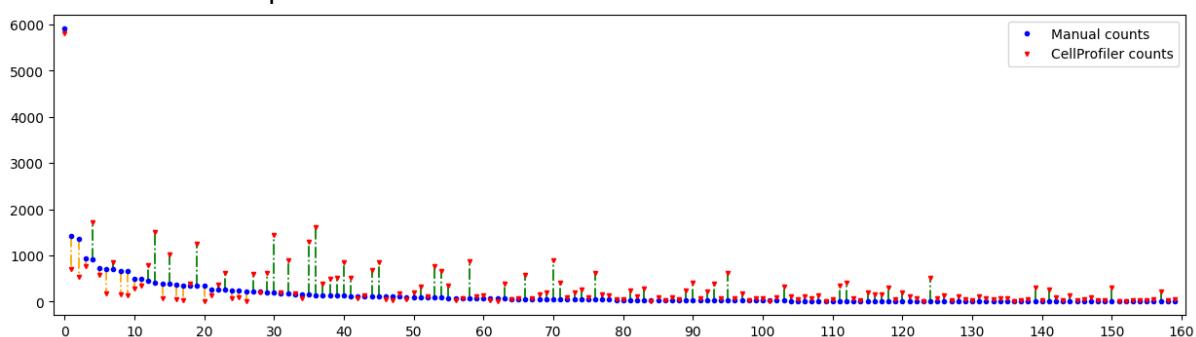
CellProfiler CD66b Pipeline Version 1.1: Correlation: 0.56



CellProfiler CD66b Pipeline Version 1.2: Correlation: 0.70

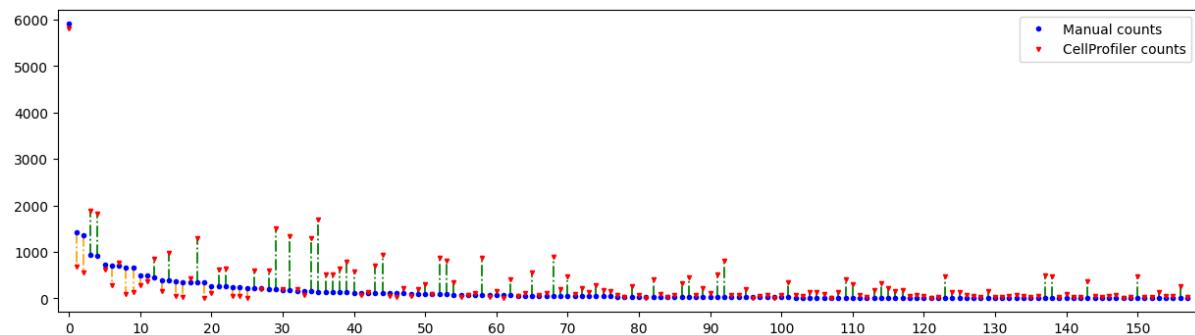


CellProfiler CD66b Pipeline Version 1.3: Correlation: 0.82

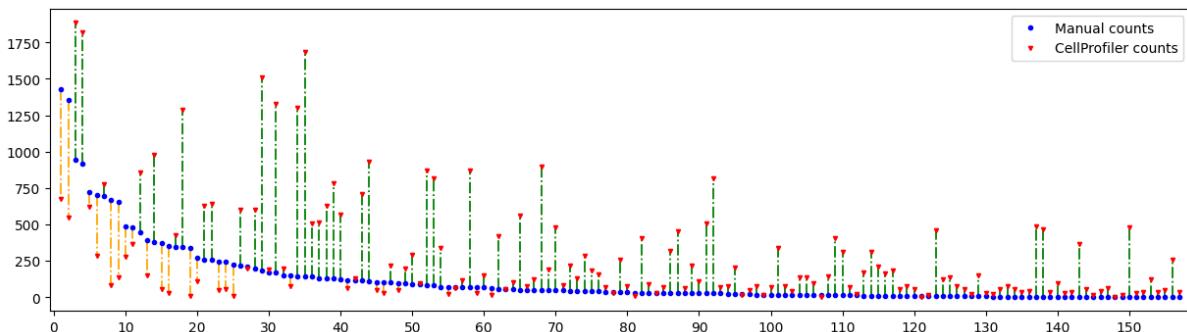


Below is the final version that will be used.

CellProfiler CD66b Pipeline Version 2.0: Correlation: 0.81

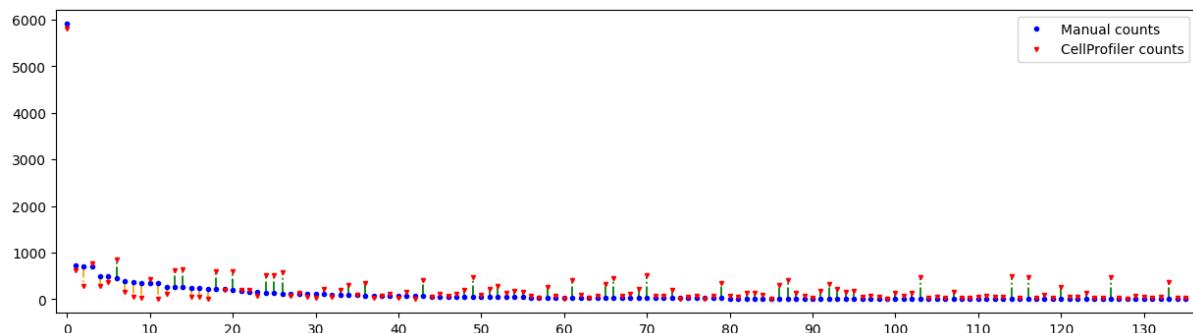


Same plot zoomed in (without the first sample):



Next, we look at the same pipeline but remove outliers by looking at the median absolute deviation with a cut-off point of 3. This means that all instances with more than 3 median absolute deviations from the median are rejected. We are left with 136 out of 158 instances.

CellProfiler CD66b Pipeline Version 2.0 (no outliers): Correlation: 0.94



As you can see the final CellProfiler pipeline contains some notable outliers. We will discuss the top 5 of these outliers below. First, we give an overview of the CellProfiler counts vs. manual counts. After that, each of these outliers will be discussed in more detail.

Cellprofiler CD66b outliers:

AE Study number	CellProfiler counts	Manual counts
AE492	1685	141
AE709	899	49
AE1182	547	1353
AE1205	817	26
AE1357	1509	187

AE492:

CellProfiler count	Manual count
1685	141

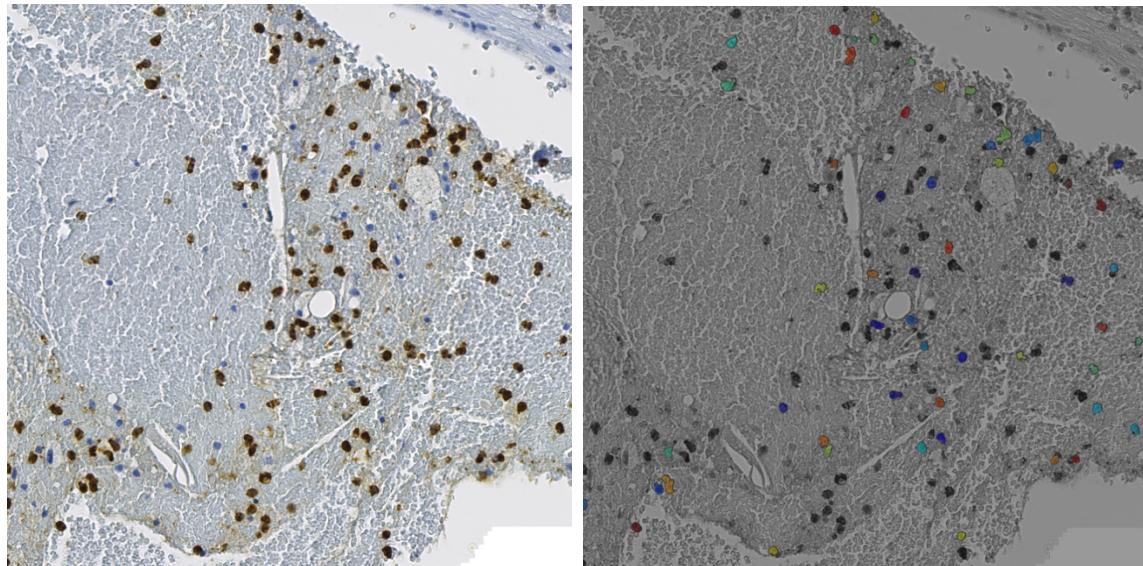
As you can see in the top 3 example shown below, the total of only these samples adds up to 131 neutrophils, which is almost the same as the total of manual count. Because there are more tiles with neutrophil cells, the manual count is incorrect. Therefore, we assume that the CellProfiler count is more accurate too the true number of neutrophils in this slide.

Top 10 tile with neutrophils

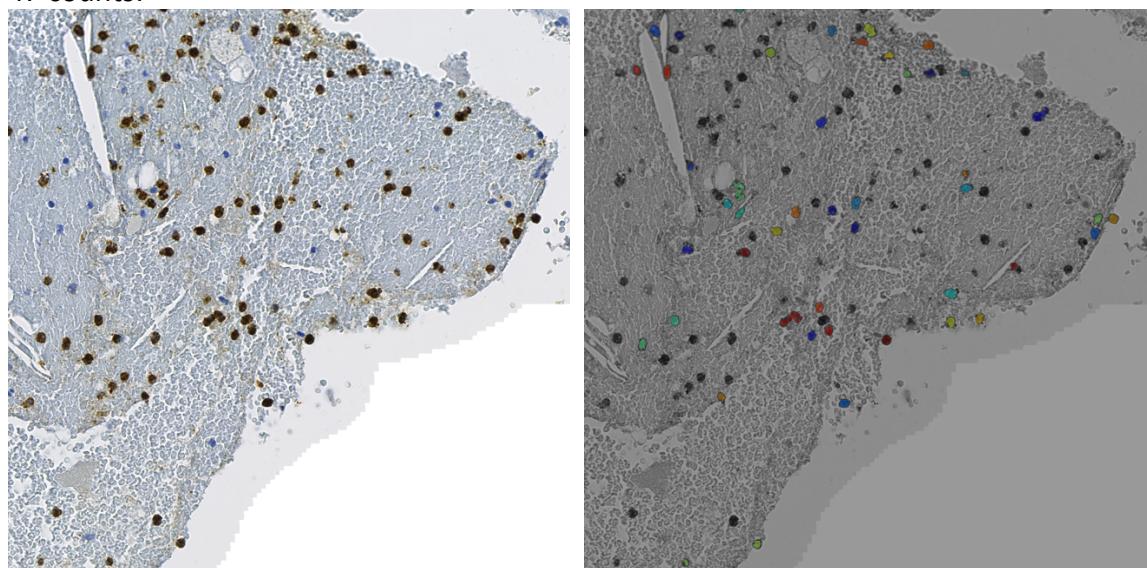
54
47
30
30
26
26
25
22
21
21

Top 3 tiles with most neutrophil counts:

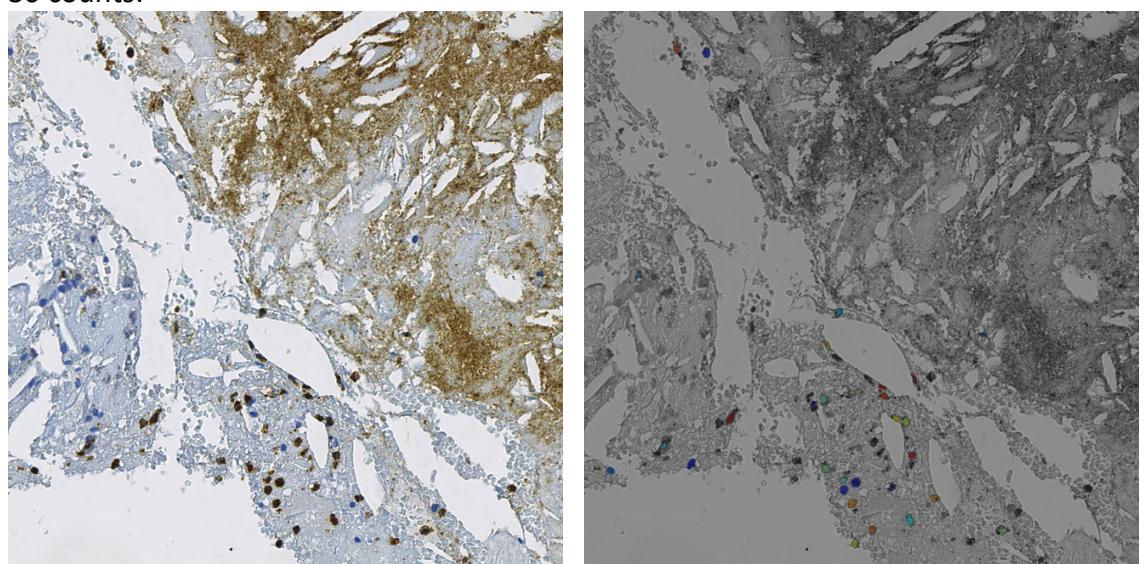
54 counts:



47 counts:



30 counts:



AE709:

CellProfiler count	Manual count
899	49

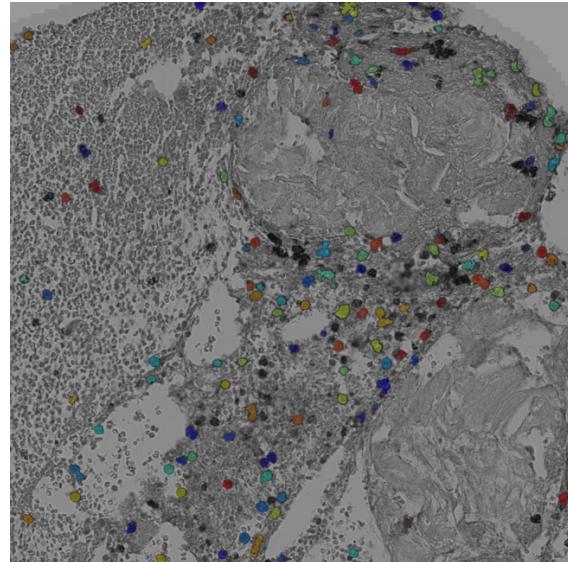
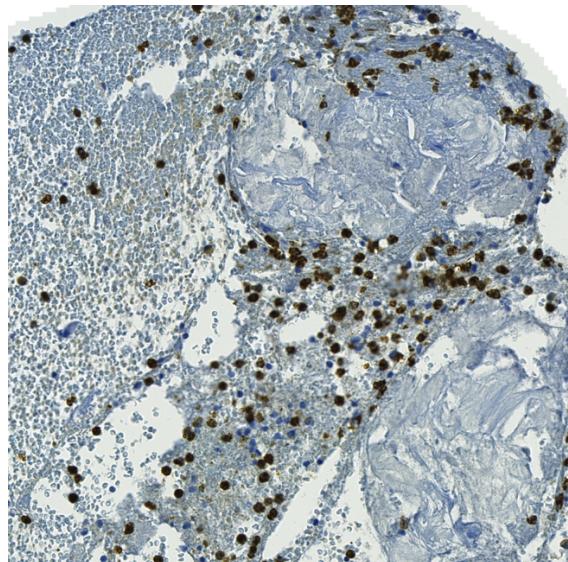
As you can see in the first example shown below, the total number of neutrophils within the first sample is more than the total manual count, which means that the manual count is incorrect. Because there are more tiles with neutrophil cells, we assume that the CellProfiler counts are more accurate too the true number of neutrophils in this slide.

Top 10 tile with neutrophils

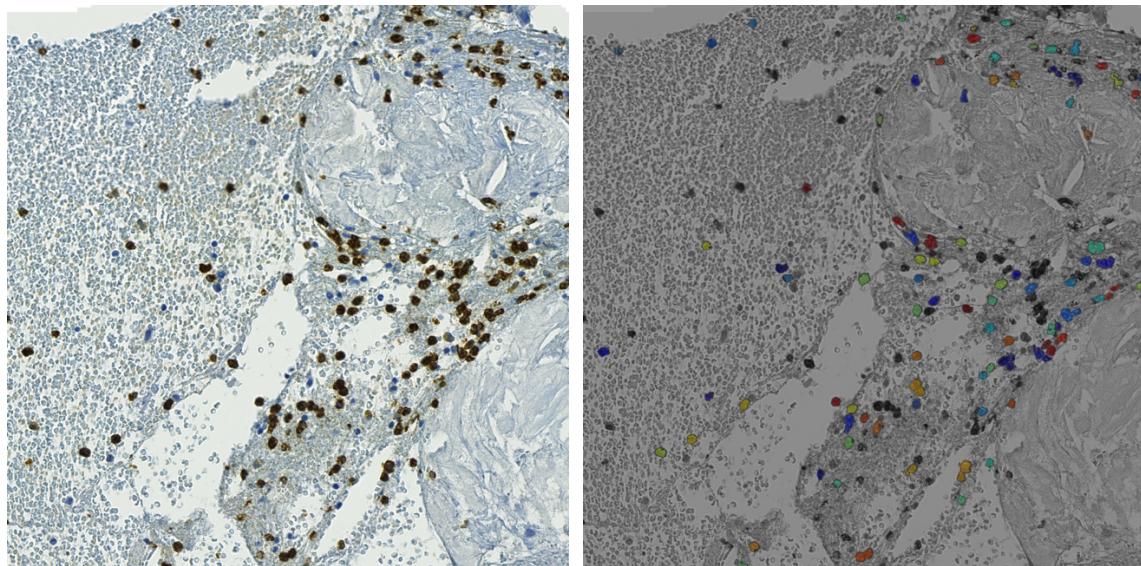
133
90
46
31
29
23
21
20
20
16

Top 3 tiles with most neutrophil counts:

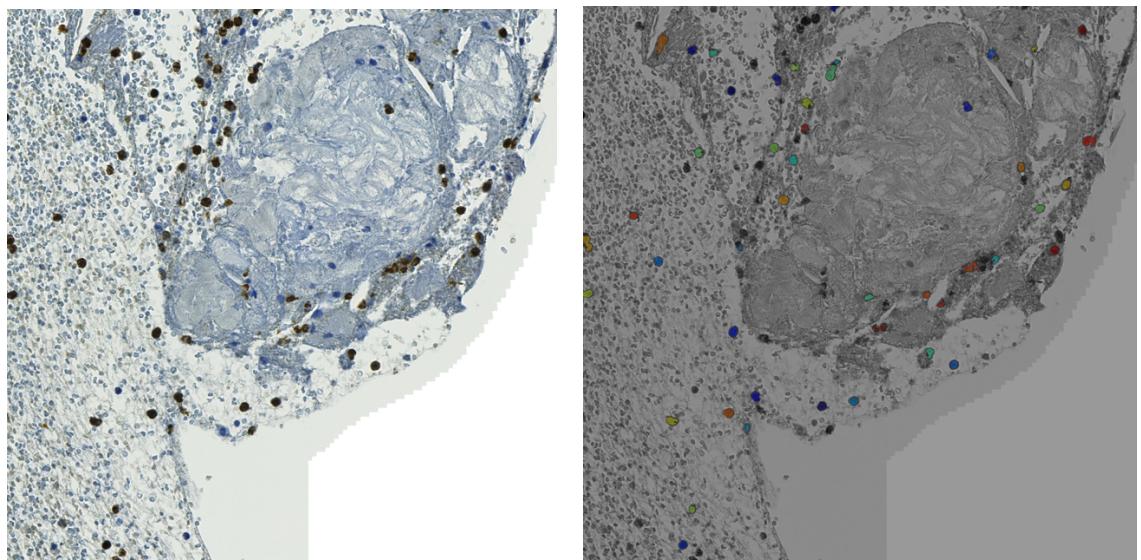
133 counts:



90 counts:



46 counts:



AE1182:

CellProfiler count	Manual count
547	1353

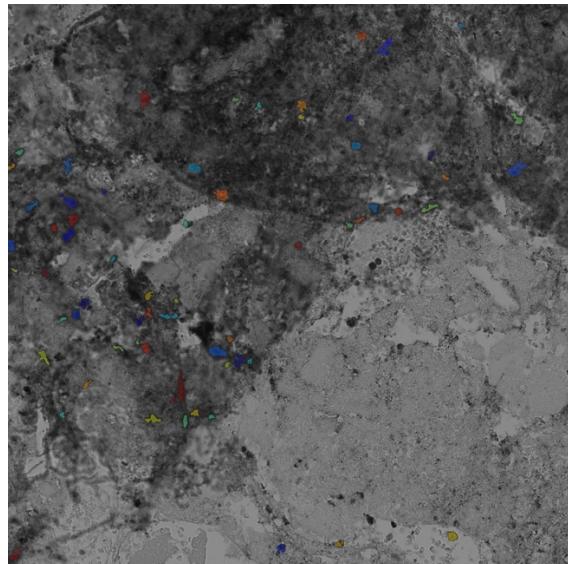
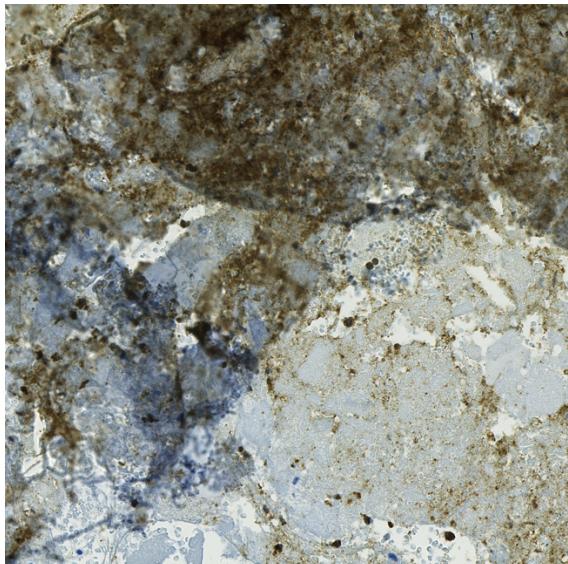
This slide is difficult to verify, but as you can see CellProfiler finds the most neutrophil counts in a slide with a lot of brown staining (which is not necessarily a neutrophil cell). CellProfiler might even count too many cells here. Therefore, the manual count cannot be accurate, because it is even higher than that of CellProfiler.

Top 10 tile with neutrophils

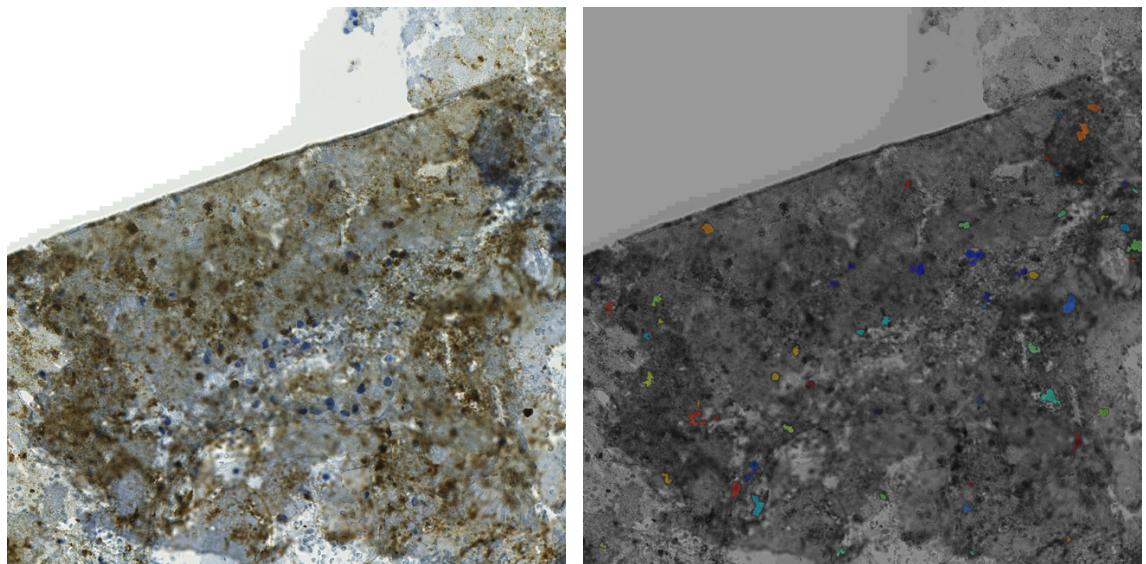
66
58
29
27
21
20
20
18
17
16

Top 3 tiles with most neutrophil counts:

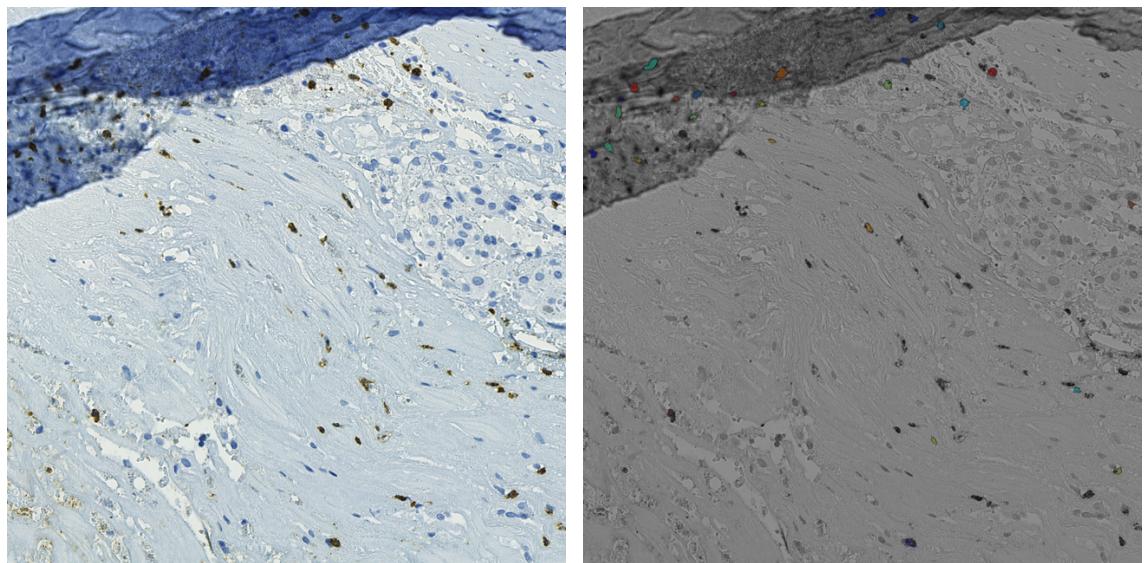
66 counts:



58 counts:



29 counts:



AE1205:

CellProfiler count	Manual count
817	26

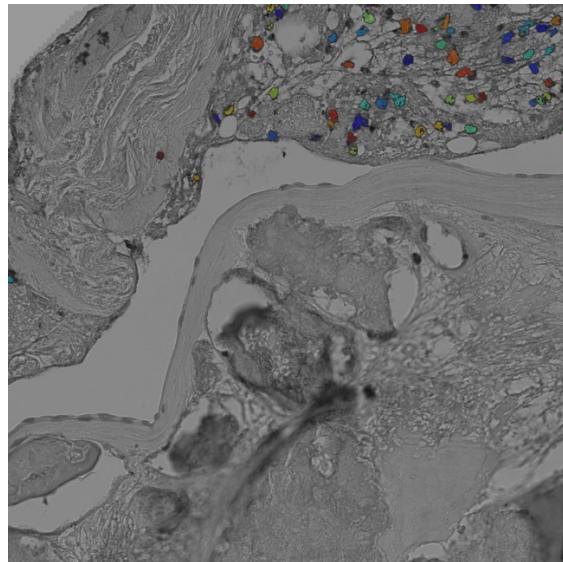
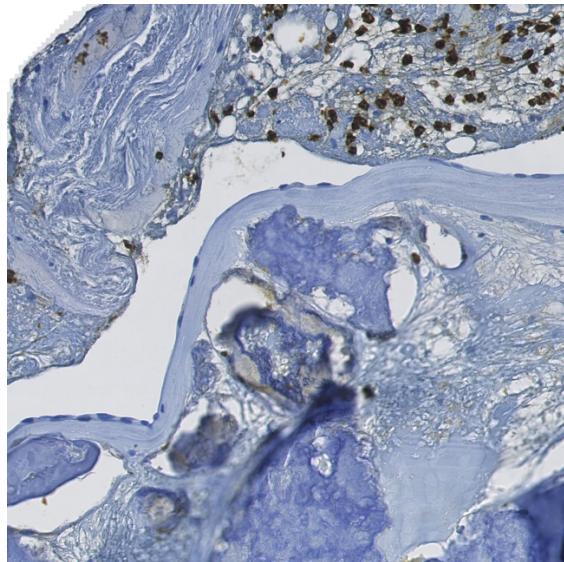
As you can see in the first example shown below, the total number of neutrophils within the first sample is more than the total manual count, which means that the manual count is incorrect. Because there are more tiles with neutrophil cells, we assume that the CellProfiler counts are more accurate too the true number of neutrophils in this slide.

Top 10 tile with neutrophils

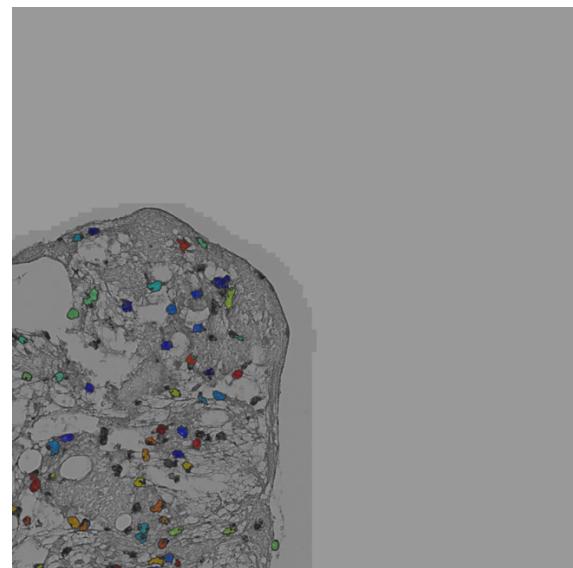
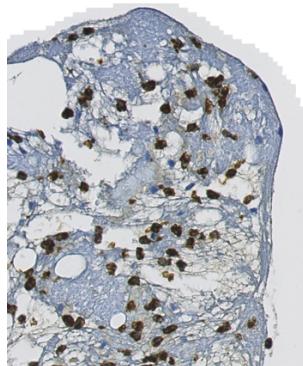
58
56
50
30
26
24
23
22
18
17

Top 3 tiles with most neutrophil counts:

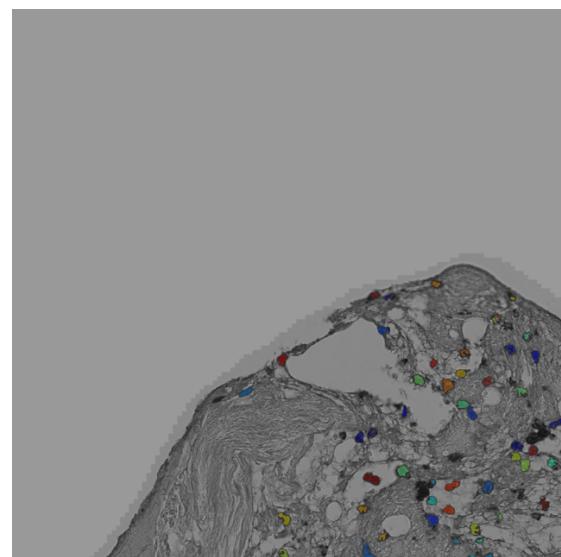
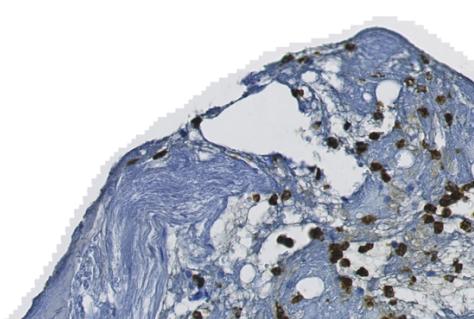
58 counts:



56 counts:



50 counts:



AE1357:

CellProfiler count	Manual count
1509	187

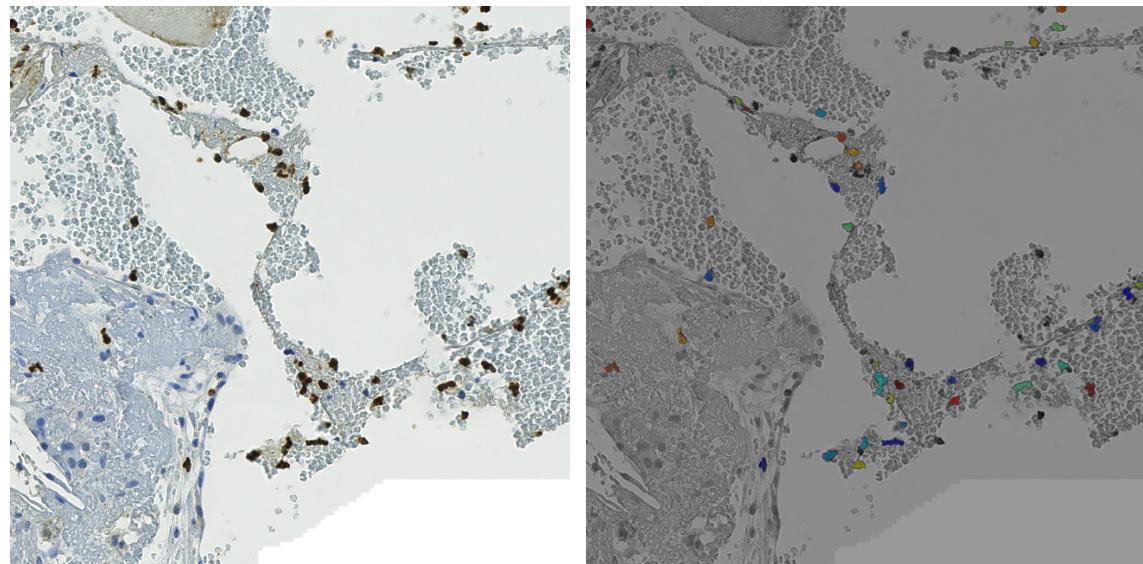
This slide has a lot of tiles with neutrophil counts around 30. As you can see from the first 3 examples these counts are quite accurate. The total sum of this would therefore be more than 187. Because of this, we assume that the CellProfiler counts are more accurate too the true number of neutrophils in this slide.

Top 10 tile with neutrophils

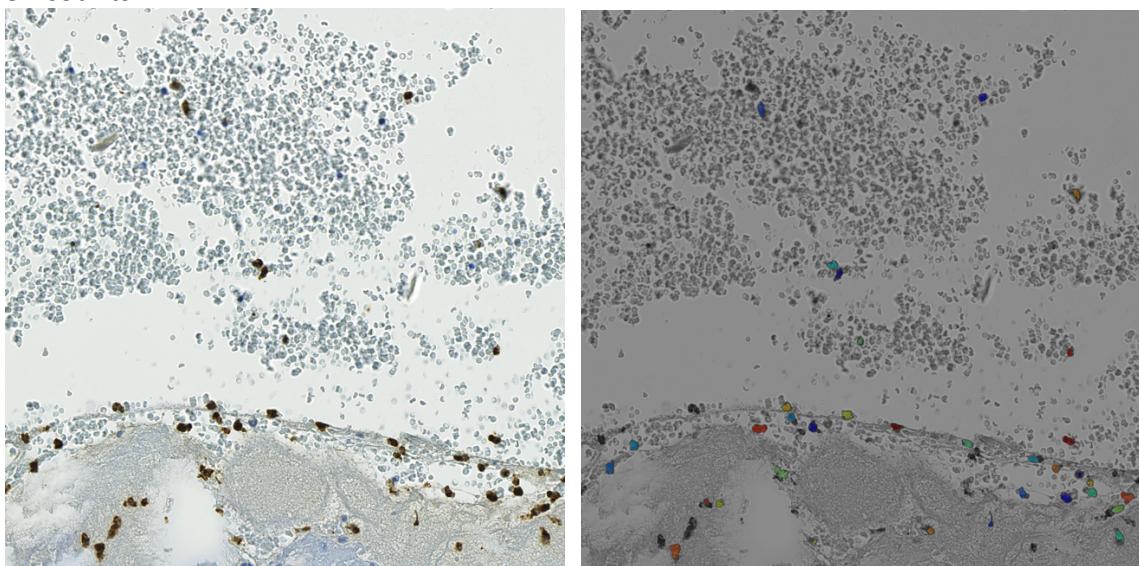
39
34
34
32
31
29
28
28
27
27

Top 3 tiles with most neutrophil counts:

39 counts:



34 counts:



34 counts:

