PHOTO BY JOHN RENNISON, THE HAMILTON SPECTATOR

Janice McFadyen at Bob Kemp hospice. She beat breast cancer once and had eight disease-free years, then it metastasized to her bones, liver and lungs.

HOW WE DID IT

Numbers are messy when it comes to science, particularly medical research

THE NUMBERS collected from experiments and clinical trials don't line up nicely in neat rows, despite the impression you're left with when you read about the latest cancer breakthrough.

Luckily, there are sophisticated mathematical tools that help organize this messiness.

Statistics are a way to try to bring some order to the numbers — to allow apples to be compared to apples, to show trends, to bring clarity to determining what's a probable explanation and what's not.

That's especially important when it comes to cancer because the numbers can be very messy for a variety of reasons.

For one, age is an important factor. Cancer, despite what you might think, skews heavily to elderly people.

Cancer is also related to genetics, diet, the environment, general health as well as certain behaviours, such as smoking or excessive drinking.

Combine all of that and it can be difficult to tease out how and why cancer developed.

A couple of things we like to measure are how many people are getting cancer and how many people are dying from it so we can figure out ways to reduce the numbers in both categories.

But that means the rates being used have to give apples-to-apples comparisons from place to place and from year to year so we know what's making a difference.

The simplest way would be to just take the number of people getting cancer — or dying from cancer — divided by the total population.

But that could be misleading, depending on the makeup of the population, because age is such an important factor with cancer.

How important is it? If you look at all cancers combined, almost 90 per cent of new cases and 95 per cent of cancer deaths occur in people 50 years of age and older.

So a community that has lots of young adults and children will have far fewer cancers than a neighbourhood with a lot of 70- and 80-year-olds.

To level the playing field, cancer specialists use a statistical tool called age standardization to allow for uniform comparisons.

First step is to pick a standard. For Canadian cancer statistics, the standard is the age makeup of the national population from 1991 — the proportion of people who were between the ages of 30 to 34, for example, and 45 to 49, and 70 to 74 and so on.

Then the population being examined — Ontario, say, or Hamilton — is adjusted so that its proportions for each age bracket match those of the standard.

The numbers of observed events, such as new cancer cases or deaths, are then adjusted to reflect the number of events that would be expected if the population being measured had the same age proportions as the 1991 standard. A key piece of information, though, is knowing the ages for each of the new cancer cases and deaths.

Using this age-standardization method, you can compare one place to another, or even look at how one place is performing from one year to another.

CANCER: A CODE RED PROJECT is based on data provided to The Spectator from Cancer Care Ontario, a provincial government agency.

The data includes the numbers of new cancer cases and cancer deaths between 2000 and 2009 for the amalgamated City of Hamilton, as well as Burlington, Grimsby, Brantford, St. Catharines, Niagara Falls, Welland, Port Colborne and Fort Erie.

The data is broken down for men and women, as well as four specific cancers: lung, breast, prostate and colorectal.

The Spectator also received cancer screening data for the same municipalities for breast, cervical and colorectal cancer for 2009, which was the most recent available year.

All of the data is broken down to the level of small geographic areas called census tracts, which are used by Statistics Canada to measure a wide variety of social, economic and health variables within urban regions of Canada.

There were 135 census tracts in the amalgamated City of Hamilton based on the previous census of 2006, and they match up well with the city's traditional neighbourhoods and former suburbs. Because of data suppression and small population sizes, five Hamilton census tracts could not be used for mapping purposes.

The Spectator was unable to carry out age standardization of rates because privacy restrictions around health data prevented the release of the specific ages for each of the new cases and deaths.

We were able to use an alternate method to create age-related rates for all of the census tracts in Hamilton, Burlington, Brantford and the major centres in Niagara.

We created rates by taking the total number of cancer cases or cancer deaths for the 10 years between 2000 and 2009 and then dividing them into the number of people 45 years of age and older in an area. We then converted that figure into a rate per 1,000 people 45 and older over the 10 years from 2000 to 2009.

This method strips out all people in a census tract or community who are younger than 45 and thus have a very low statistical chance of getting cancer or dying from it.

We used this same method to also create the same rates for Canada, to use for comparison purposes.

IT'S NOT PERFECT, we acknowledge, and there are limitations to our methodology.

The populations of census tracts would have fluctuated between 2000 and 2009, and it's possible that some people might have changed census tracts or even municipalities between the date of the 2006 census and the time they were diagnosed with cancer or died from it. The total number of cancer cases and deaths used to create the rates includes a small fraction that would have occurred in people younger than 45 years of age.

Cancer incidence, mortality and screening rates were then compared with selected social and economic variables — such as income, poverty levels and education — that are tracked by Statistics Canada. For this project, we used social and economic data from the 2006 census, which fell almost at the halfway point of the 2000 to 2009 time period of the cancer data we received.

The cancer and socioeconomic data sets were then forwarded to Patrick DeLuca, a mapping and statistics expert in McMaster University's Centre for Spatial Analysis. He agreed to collaborate on a pro bono basis.

DeLuca used the data to create a number of maps.

Some of the maps are rate-based.

Some of the maps are quintile maps, which distribute the 130 census tracts into five quintiles. Each quintile represents 20 per cent of the 130 census tracts, from highest to lowest for the variable being measured.

The maps, rankings and rates were then analyzed to examine the connections that exist between cancer and socioeconomic variables in Hamilton's neighbourhoods.

ALL NEW CANCER CASES 2000-2009