The data provided here relate to the paper Goudge et al “Can lay health workers support the management of hypertension? Findings of a cluster randomised trial in South Africa” BMJ Global Health 2018;3: e000577 doi: 10.1136/bmjgh-2017-000577

Data provided consists of:

1. A Stata data file: nkateko trial final data.dta
2. Two pdf files with the before and after survey questionnaires:
	1. Nkateko-Surveillance-V6-20130927.pdf (first survey)
	2. Nkateko-Surveillance-V10-20150819.pdf (second survey)

The samples for both surveys were drawn from the census database maintained by the Agincourt Health and Demographic Surveillance System (<http://www.agincourt.co.za/>) and demographic data were drawn down from the census and added to sample file. These variables are usually indicated by the variable label. Apart from a few created variables such as age and age group, all the other data come from the two surveys and the variable names indicate which question in the survey the data comes from. The variable “survey\_round” indicates which survey information is drawn from. Because there were two independent samples there are a few hundred individuals that appear in both surveys.

During the first survey there was problem with a field worker who turned out to be making up data. All the work she had done had to be discarded and this is indicated in the variable name “dont\_analyze”.

In the research area, there was no obligation for people to use a particular clinic. There were eight clinics in the trial but some individuals used clinics outside the trial area, and were not included in the outcome analysis.

Margaret Thorogood

m.thorogood@warwick.ac.uk