

# Determining disease prevalence and expected effects of interventions for sample size calculations in general practice research: An example



National Prescribing Service Limited

HK Parsons, F Garden, M Williamson, JM Mackson  
National Prescribing Service Limited

## Introduction

General practice is a complex environment in which to evaluate the impact of interventions on prescribing behaviour using data from clinical software. It is not always possible to extract relevant data from clinical software, there is much variability in the way GPs record clinical data and uncoded diagnoses are frequently used.

This project aimed to understand the opportunities for changing GP prescribing behaviour, based on disease prevalence in the practice and patient visits. The project also aimed to estimate the number of practices (sample size) needed to show a significant change from a proposed intervention.

This multi-faceted intervention would include extraction and feedback of prescribing data from clinical software to GPs, facilitated peer group discussion and education and continuous quality improvement cycles. A cluster randomised controlled trial (RCT) is planned to test the effectiveness of the intervention in changing prescribing.

## Methods

To understand the opportunities that GPs have to change prescribing, we identified two areas and two corresponding indicators where improved prescribing has led to improved outcomes for patients: increased use of beta blockers in patients with heart failure and intensifying antihypertensive therapy in hypertensive patients not at target blood pressure.

We used two data sources to estimate the prevalence of these conditions in general practice: Supplementary Analysis of Nominated Data (SAND) studies<sup>1,2</sup>, and data recorded in GP clinical software from the General Practice Research Network (GPRN). We also used the GPRN data to estimate the number of patient visits in a 6 month period.

When analysing GPRN data, we used data from January 2004 to January 2008 for practices that had been involved in GPRN for more than 23 months. Only diagnoses that had been entered into the clinical software as codes (i.e. not free text) were used as this reflects the data available for automated data extraction from clinical software.

We searched the literature and NPS data for baseline prescribing rates and the likely effect sizes of implementing a similar intervention. We considered the prevalence, frequency of visits and expected effect sizes to calculate sample sizes and feasibility for the RCT of the intervention.

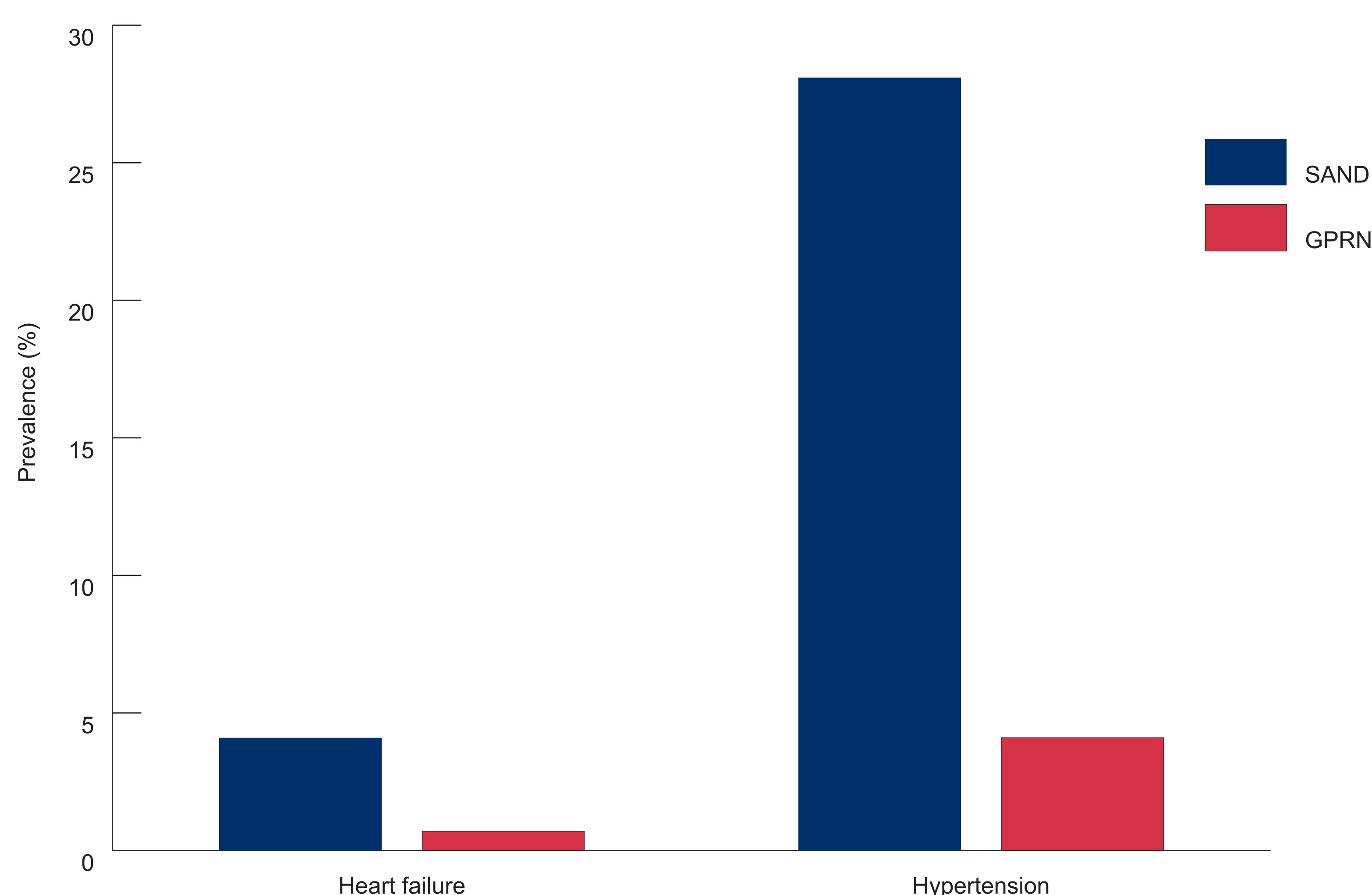
Sample sizes were calculated with an assumed intracluster correlation of 0.1 at 80% power and 5% significance level using the University of Aberdeen sample size calculator, which adjusts for the clustering of GPs and patients in practices (version 1.0.2, available at [www.abdn.ac.uk/hsru/epp/cluster.shtml](http://www.abdn.ac.uk/hsru/epp/cluster.shtml)).<sup>3</sup>

## Results

### Disease prevalence

Estimates of the prevalence for heart failure and hypertension differed markedly between SAND and GPRN data (see Figure 1).

Figure 1. Disease prevalence in two general practice data sets



## Estimated effects of intervention

As GPRN data reflects the prescribing data that will be available for use in the intervention, this prevalence was used. It was assumed that an average practice of 3 GPs would have at least 3000 active, living patients. GPRN data showed that in a 6 month period, approximately 50% of patients with chronic diseases such as heart failure and hypertension are seen by a GP. Numbers of these patients expected to visit a practice in a 6 month study period are presented in Table 1. These represent opportunities for the GP to review management and make changes to prescribing.

Table 1 also shows the expected effects of the proposed intervention and the number of practices required to show these effects.

Table 1. Prevalence, expected effect and sample size for each diagnosis

	Heart failure*	Hypertension
Prevalence (GPRN data)	0.7%	4.1%
No. patients visiting practice in a 6 month period (GPRN data)	10	49
Expected % pre-intervention	40% of patients use a beta blocker <sup>4</sup>	60% of patients using an antihypertensive have BP above target <sup>5</sup>
Expected effect of intervention	15% absolute increase in beta blocker use <sup>4</sup>	10% absolute decrease in patients above target BP <sup>6</sup>
Sample size required	66 practices	92 practices

\* Includes patients using a loop diuretic with an angiotensin converting enzyme (ACE) inhibitor or angiotensin II-receptor antagonist.

## Discussion

- SAND data collection requires GPs to record a diagnosis for each patient encounter, whereas GPRN data relies on diagnoses recorded within clinical software. Although SAND data may more accurately reflect disease prevalence in general practice, GPRN data provided a better estimate of the number of patients available for the GP to review as part of the intervention.
- The differences in prevalence estimates between GPRN data and SAND data were most likely due to poor recording of diagnoses in clinical software used in GPRN.
- A time period longer than 6 months may be required to obtain sufficient numbers of patients for an RCT of the proposed intervention.
- Taking into account likely dropouts, over 100 practices will be required to show the expected improvements in beta blocker prescribing rates and in achievement of BP targets.

## Implications for future research

- Researchers should be aware that the actual disease prevalence may not be reflected in the data available from clinical software systems, particularly when using data extraction tools.
- GPs participating in research using data extracted from clinical software should use coded diagnoses so that maximal numbers of patients can be identified for an intervention.
- Researchers may benefit from using GPRN data for sample size calculation and implementation planning when using data from clinical software.

## References

1. AIHW Australian GP Statistics and Classification Centre, 2006. SAND abstract No. 90 from the BEACH program: Prevalence, management and investigations for chronic heart failure in general practice patients. Sydney: AGPSCC University of Sydney.
2. AIHW Australian GP Statistics and Classification Centre, 2005. SAND abstract No. 79 from the BEACH program: Hypertension and dyslipidaemia – comorbidity and management in general practice patients. Sydney: AGPSCC University of Sydney.
3. Campbell MK, Thomson S, Ramsey CR, et al. Sample size calculator for cluster randomised trials. *Comput Biol Med* 2004;34(2):113–125.
4. Hickey A, Scott I, Denaro C, et al. Using clinical indicators in a quality improvement programme targeting cardiac care. *Int J Qual Health Care* 2004;16 (suppl 1): i11–i25.
5. National Institute of Clinical Studies, 2005. Evidence-practice gaps report, Volume 2. National Institute of Clinical Studies: Melbourne, 2005.
6. Walsh JME, McDonald KM, Shojania KG, et al. Quality improvement strategies for hypertension management: a systematic review. *Med. Care* 2006; 44:646–57.

Data sources include: General Practice Research Network – Health Communication Network.