INTRODUCTION
Over 300 million people suffer from asthma globally, which makes it one of the most common chronic diseases in the world.1,2 International guidelines such as the Global Initiative for Asthma (GINA) formulate the long-term goals of asthma treatment as the accomplishment of asthma control and reducing the exacerbation rate.3–6 Although people with asthma generally perceive their asthma control positively, the level of actual clinical asthma control is, despite its step-wise approach, more of a one-size-fits-all approach, that is not personalised according to individual patient characteristics. Tailored interventions using FeNO7 or sputum eosinophils8, for example have proven successful in reducing steroid dose and exacerbation rate, respectively.

Currently, several different asthma phenotypes that cluster specific patient characteristics have been described.9–13 Managing asthma patients according to these phenotypes might be an approach that allows for individualising treatment. Thus far most of the studies identifying phenotypes have been performed in hospital care. However, in many countries, the majority of the individuals with asthma are treated in a primary care setting.12–16 Therefore, it is important to assess whether these phenotypes are similar in primary care. Additionally, the majority of the current hospital care-based phenotypes require certain measurements not easily available in primary care. This hampers implementation and it would be preferable if primary care phenotypes consist of measurements easily obtainable to the general practitioner (GP) or practice nurse (PN). Therefore, as the first part of our study we aimed to cluster primary care asthma patients into distinct phenotypes based on easily obtainable patient characteristics using a clustering strategy.

Furthermore, continuing onward with these phenotypes, efficiency of asthma management might be improved if it is known which phenotypes require more frequent assessments by a GP or PN, because they present a greater risk of future adverse outcomes and which phenotypes could be safely assessed less frequently due to more favourable long-term asthma outcomes. Therefore, in the second part of our study we determined the long-term differences in important asthma outcomes between different phenotypes. We aimed to assess differences in our established phenotypes in terms of asthma control, quality of life, medication usage and exacerbation frequency, after a 12-month follow-up period.

RESULTS
The results described here were generated in a five step process, which is described in detail in our methods section. In the first step we assessed potential variables to be included in our cluster analysis and these are listed in Table 1. Overall, our population is comparable to the general asthma population in primary care in Europe, including a higher percentage of females (Table 1).15–17 Next, we performed a factor analysis which showed four distinct factors. These factors were defined by: (1) the level of self-reported patient health, labelled 'functioning'; (2) the combination of...
prescribed medication, labelled ‘medication’; (3) the atopic status, labelled ‘allergy’ and (4) the socioeconomic status, labelled ‘SES’. The results of the factor analysis are reported in Table 2, results are standardised z-scores and positive scores represent an above average result (see methods). Subsequently in step 3 data transformation was performed for selected variables. FEV1 was adjusted for gender and missing variables were imputed. Then we performed a cluster analysis with 14 variables (including factors), all variables were continuous except gender.

An agglomerative hierarchical cluster analysis with baseline data showed that the optimal amount of clusters was five. A general overview of the clusters is presented in Table 2. In summary, the first phenotype consisted of patients with an early diagnosis of asthma and a high prevalence of allergies. This phenotype also reported the highest functioning scores and socio-economic status (SES). Most patients in this phenotype were male. For further reference, this phenotype was labelled ‘early atopic’.

The second phenotype, labelled ‘late onset female’, was the largest group. It consisted mainly of females and the mean age of diagnosis was slightly higher than in the other phenotypes.

The third phenotype, labelled ‘reversible’, consisted of patients with high reversibility after bronchodilation and the lowest baseline FEV1 percentage predicted. This group also reported the most symptoms in the Asthma Control Questionnaire (ACQ6). The fourth phenotype, labelled ‘smokers’, included patients with the highest amount of pack years smoked. This not necessarily meant the patients were still active smokers during the study (a total of 48% was actively smoking). This group also reported the lowest SES.

The fifth and last phenotype, labelled ‘exacerbators’, comprised of patients who frequently experienced exacerbations in the year previous to the study. They also reported relatively high ACQ scores.

Long-term outcomes
Finally we analysed long-term outcomes of important asthma parameters for each of the phenotypes.

Table 1. Baseline data of total research population

| Variable                              | Total (n = 611) |
|---------------------------------------|----------------|
| Gender (% female)                     | 68.4           |
| Age, yr (SD)                          | 39.4 (9.1)     |
| Age of diagnose, yr (SD)              | 20.8 (14.4)    |
| Length, cm (SD)                       | 172.4 (10.1)   |
| Weight, kg (SD)                       | 79.0 (17.7)    |
| BMI, kg/m² (SD)                       | 26.4 (5.3)     |
| ACQ6a                                 | 1.0            |
| AQLOQ                                 | 5.7            |
| Reversibility, %a,b                    | 6.6            |
| Active smokers (% patients)           | 14             |
| Pack years smoked, yr                 | 4.7            |
| Exacerbations in past 12 months, no per patienta | 0.67          |
| FeNOa                                 | 23.3           |
| FEV1 (% predicted)a                   | 95             |

Table 2. Clustering results

| Variable                      | Cluster 1 (n = 170) | Cluster 2 (n = 236) | Cluster 3 (n = 62) | Cluster 4 (n = 62) | Cluster 5 (n = 81) |
|-------------------------------|---------------------|---------------------|--------------------|--------------------|--------------------|
| Gender (% female)             | 45.3                | 85.2                | 75.8               | 51.6               | 75.3               |
| Age, yr                       | 34.8                | 42.5                | 34.5               | 44.5               | 39.7               |
| Age of diagnose, yr           | 13.7                | 27.1                | 16.5               | 23.5               | 18.5               |
| BMI, kg/m²                    | 25.2                | 27.4                | 25.1               | 25.5               | 27.5               |
| ACQ6                          | 0.5                 | 0.9                 | 1.7                | 1.0                | 1.6                |
| AQLOQ                         | 6.3                 | 5.6                 | 5.3                | 5.8                | 5.1                |
| Reversibility, %              | -1.8                | 2.7                 | 35.1               | 10.7               | 10.9               |
| Pack years smoked, yr         | 1.2                 | 2.9                 | 2.1                | 24.0               | 4.4                |
| Exacerbations in past 12 months, no per patient | 0.1                | 0.4                 | 0.5                | 0.4                | 3.0                |
| FEV1 (% predicted)            | 96                  | 100                 | 70                 | 89                 | 99                 |

Factorsa

| SES                           | 0.20                | 0.01                | -0.16              | -0.48              | 0.03               |
| Allergy                       | 0.27                | -0.20               | 0.05               | -0.33              | 0.22               |
| Functioning                   | 0.42                | -0.18               | -0.16              | 0.15               | -0.37              |
| Medication                    | -0.38               | 0.18                | 0.05               | 0.01               | 0.24               |

| Factorsa                      | Factor scores are standardised Z-scores with mean = 0 and standard deviation = 1 |
Asthma quality of life questionnaire (AQLQ). Over the 12-month follow-up, all phenotypes gradually improve in quality of life over time, with the ‘exacerbators’ phenotype showing the most improvement. After 12-month follow-up the ‘early atopic’ phenotype has the best asthma related quality of life (AQLQ = 6.35). This was >0.5 points above both the ‘reversible’ (5.53) and the ‘exacerbators’ (5.49) phenotype, again indicating a clinically meaningful MID for the AQLQ. The AQLQ-results of the different phenotypes never cross each other.

Exacerbations. In Fig. 2 the cumulative mean amount of exacerbations during the study is presented. The ‘exacerbators’ phenotype experienced the highest absolute amount of exacerbations in the follow-up period. The ‘smoker’ phenotype also had a relative high amount of exacerbations compared to the other phenotypes.

Medication usage. Figure 3a shows the total dosage of inhaled corticosteroid (ICS) after 12-month follow-up for each of the phenotypes and Fig. 3b shows the change during the study in guideline derived treatment steps. The ‘exacerbators’ phenotype were prescribed the highest ICS dosage after 12 months and also increased their dosage the most during the study. The ‘early atopic’ phenotype were prescribed the lowest dosage from start to finish.

DISCUSSION

Main findings

In this study we established five distinct asthma phenotypes in primary care. Important long-term asthma outcomes, such as asthma control, quality of life, exacerbation rate and medication usage differed between these phenotypes after 12-month follow-up. Therefore, taking into account asthma phenotypes allows for a more personalised treatment approach and in so doing can have important consequences for long-term asthma outcomes.

Interpretation of findings in relation to previously published work

In our study in primary care we established five distinct asthma phenotypes. Previously, Haldar et al. found three clusters in a primary care population, namely: ‘early-onset atopic asthma’, ‘obese noneosinophilic’ and ‘benign asthma’. In our study the ‘early-onset atopic’ phenotype had the most favourable outcomes, in terms of asthma control, quality of life and exacerbations, while using the lowest dosage of medication. This is in contrast to previous studies that suggest that early onset atopic patients use higher ICS dosage than other phenotypes. This can be attributed to our additional specific ‘exacerbators’ phenotype, which can be seen as the subset of patients with the worst long-term asthma outcomes. Therefore, by using our phenotypes, patients with the worst long-term outcomes are separated from the other phenotypes and consequently treatment can be more targeted to those with an increased risk.

The clusters in the study by Ortega showed, despite the high average body mass index (BMI), some familiarities with our clusters. Cluster 3 has relative many allergic patients but since this is not the group with the youngest patients, this does not directly correspond with our ‘early onset atopic’ phenotype. Cluster 7 could correspond with our ‘exacerbators’ phenotype and cluster 5 could correspond with our ‘smokers’ phenotype. However, acknowledging the differences in variables used, the clusters are not directly the same.

An important reason that we found several different phenotypes from previous studies was that we added variables easily assessed in primary care. We also included FeNO as a variable in our analyses as it is now available in primary care and is shown to be cost-effective. We also included the variable pack years smoked in the cluster analysis. The amount of pack years smoked highly influences respiratory disease progression. This is
somewhat reflected in the ‘smokers’ phenotype with a higher exacerbation rate. In contrast to previous studies, we did not detect one specific cluster of obese patients.\textsuperscript{11,20} One possible explanation could be the relatively high mean BMI (26.4) in our population, which allows for less distinction purely based on weight. This high mean BMI is comparable with the average BMI in western society.\textsuperscript{23–25}

Strengths and limitations

An important strength of our study is that we specifically studied an asthma population managed in primary care, while most former studies assessed phenotypes in hospital care. In our study we included many patients ($n=611$) and we had 12-month follow-up data. We specifically used variables that were easily measurable and did not look at eosinophilic airway inflammation and comparable laboratory-requiring variables.

Another strength of our study is the heterogeneous group of primary care patients, of which many different variables were collected during the original trial, which allowed us to use many variables for clustering. The variable selection and analyses were based on results from previous literature.

A limitation of our study was that the differences within a phenotype (from baseline to 12 months) were usually higher than differences between phenotypes (at a certain point). While the ‘exacerbators’ phenotype had the relative highest risk of developing an exacerbation in the future, the absolute amount of exacerbations they experienced declined in the study year (mean 0.85 exacerbations per patient per year (exc/pt/yr)), compared to the year prior to the study (3.0 exc/pt/yr). This can partly be attributed to the beneficial effect of participating in a clinical trial, since the total amount of exacerbations in the whole research population also decreased (0.67 to 0.24 exc/pt/yr).

Another limitation of our study was the amount of patients per phenotype. Even though our general asthma population comprised of many patients, some phenotypes were too small (smallest phenotype $n=62$) to perform further analysis, which resulted in our study having a more descriptive character. Further studies require more patients per phenotype to perform optimal treatment strategy analyses.

In our study we used the data set from the ACCURATE trial, in which patients were randomised into three different treatment strategies (‘partially controlled’, ‘controlled’ and ‘FeNO’).\textsuperscript{22,26} Our phenotypes were determined using baseline data from this trial, which were obtained before randomisation. Therefore the treatment strategies of the original trial had no influence on the establishment of the phenotypes. However, long-term outcomes could have been influenced if one of the treatment strategies of the original trial would have been overrepresented in one of our phenotypes. Therefore we performed a sub-analysis and this showed that treatment strategies were roughly equally distributed amongst the phenotypes (data not shown).

Implications

Determining different phenotypes is of aid in identifying patients at risk, enabling primary care physicians to guide management, including frequency of control visits, treatment decisions and proper allocation of patients to primary care or hospital care management. To date, to describe people with asthma, usually mean scores are used, while the spread of scores is high. This makes it hard to make general statements about asthma patients in primary care and agglomerating all these different patients together leads to ineffective use of the healthcare system. Differentiating by using phenotypes might tackle this problem and the differences in long-term outcomes between the phenotypes in our study show that they function differently from one another, which underscores the benefit of phenotyping asthma patients in primary care.

For example, the ‘exacerbators’ phenotype clearly has the least favourable outcomes, even though they are prescribed the highest level of medication. Therefore, only increasing medication dosage for these patients is insufficient. Even though exacerbation rate improved in comparison to the year before the study, it is still high, while patients might experience additional side-effects. This suggests these patients additionally require a more thorough review of other (non-pharmacological) treatment options and if no clear improvements can be made, they should probably be referred to specialised care.\textsuperscript{26} On the other side of the spectrum is the early onset atopic phenotype. This group had good results on both ACQ and AQLQ and they had the lowest medication level. Therefore we performed a sub-analysis and this showed that treatment strategies were roughly equally distributed amongst the phenotypes (data not shown).
healthcare system as a whole if visits to the GP-office for these phenotypes could be limited.

Further research should focus on the implementation of phenotype usage in primary care. Firstly it must be assessed how to easily define a particular patients’ phenotype within current asthma management and how often this should be repeated in order to continually improve long-term asthma outcomes. Also, while we did show long-term outcomes are different for phenotypes, it is currently not certain that changing therapy according to this knowledge will actually lead to improved outcomes. Although this seems likely, a future prospective study is needed, which could also assess the consequences of implementing phenotypes in general practices, since if it is time-consuming while real life impact on long-term outcomes are negligible, use of phenotypes is questionable.

CONCLUSIONS
Clustering of primary care asthma patients resulted in five distinct phenotypes. These phenotypes differ in long-term outcomes, as measured by ACQ, AQLQ, exacerbation frequency and medication usage. Using these phenotypes in asthma treatment may result in a personalisation of treatment, based on individual long-term outcomes.

METHODS
Subjects
For this study we used data collected in the Asthma Control Cost-Utility Randomised Trial (ACCURATE), a multicentre randomised trial performed in the Netherlands. The goal of this original study was to assess three different treatment strategies of asthma. The results of this trial have previously been published. For a detailed description of study procedures and participants see the published protocol.

In short, participants were selected in primary care offices in both rural and urban areas in the Netherlands. All patients gave informed consent. Inclusion criteria were a doctors diagnosis of asthma, age 18–50 year-old, and a prescription for inhalation corticosteroid’s (ICSs) for at least 3 months in the previous year. Patients were excluded if they had used an oral corticosteroid in the month prior to the study, were unable to speak Dutch or suffered from severe comorbidity. Methods were performed in accordance with relevant regulations/guidelines. The trial was approved by the Medical Ethics Committee of Leiden University Medical Center and registered at http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=1756.

Study design
The ACCURATE study was a cluster randomised trial with a follow-up of 12 months. For the establishment of the phenotypes baseline data were used, which contained data of patients before randomisation altered outcomes. Participants were reviewed every 3 months and if necessary treatment was adjusted. Additionally patients filled out online questionnaires. For our study we retrospectively analysed the data. Long-term outcomes of phenotypes were assessed using the data from the three monthly assessments and online questionnaires.

This analysis in this study consisted of the following five steps:
1. Identification of potentially relevant variables
2. Factor analysis
3. Data transformation, if necessary
4. Cluster analysis
5. Comparison of long-term outcomes

In the first step all potentially relevant variables were identified based on previous studies and relevant guidelines. General patient characteristics (gender, age, BMI, income and education level) and asthma specific variables were selected. All variables used for the analysis are listed in Table 1. To determine the different phenotypes, baseline data were used to approximate real life circumstances.

In the second step, we performed a factor analysis. A factor analysis was conducted before clustering to combine interrelated variables that measure the same construct, namely: SES, allergy, functioning scores on the sf36 questionnaire and medication. A varimax rotation was used to maximise variance.

In the third step, part of the data was transformed to increase validity of the analysis and FEV1 was adjusted for gender and extreme outliers. Data were imputed if there were missing values at random, with STATA-command imputation. Additionally all variables used for clustering were standardised, with a mean of 0 and a standard deviation of ±1 in order to give all variables equal weight in the cluster analysis.

The fourth step consisted of a Wards linkage cluster analysis. This is the most optimal agglomerative hierarchical clustering strategy for the given data. Following Ward’s criterion for minimum variance the within-cluster variance will be as small as possible for the most similar clusters. The distance between the clusters was measured with the squared Euclidean distance.

The final step of the analysis was to assess long-term outcomes for every individual phenotype.

The ACQ6, the AQLQ at 12 month-follow-up, the amount of severe exacerbations during the study-period and medication prescription were selected as long-term outcomes. The ACQ6 is a six-question questionnaire aimed to map asthma control, it scores between 0–6 and higher scores represent worse asthma control. The AQLQ assesses asthma related quality of life, the score has a range from 0 to 7 with higher scores representing a higher quality of life. A severe asthma exacerbation was defined as a course of oral prednisolone prescribed for worsening asthma for three or more days, or an emergency department visit/hospitalisation due to asthma. Medication prescription was defined as the total usage of ICSs, recalculated to beclomethasone equivalent dosage. Additionally the use of long-acting beta agonists was presented separately. Finally all medication usage was also classified into one ‘medication step’, based on the steps presented in the NAEPP guidelines.

Data availability
Data analysed in this study are available upon request.

ACKNOWLEDGEMENTS
The study was performed in association with the department of medical decision making of the Leiden University Medical Centre.

AUTHOR CONTRIBUTIONS
The main author and data analyst of this article was R.K. P.H. and J.S. were important co-authors and reviewed analyses. Data for the original trial were collected by P.H., R. L., J.S.-S., P.A. and T. S. All authors were involved in critically revising the manuscript and approved the final version.

ADDITIONAL INFORMATION
Competing interests: The authors declare that they have no competing financial interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES
1. Masoli, M., Fabian, D., Holt, S. & Beasley, R., Global Initiative for Asthma. P. The global burden of asthma: executive summary of the GINA dissemination committee report. Allergy 59, 469–478 (2004).
2. GINA. Global Strategy for Asthma Management and Prevention http://www.ginasthma.org/ (2016).
3. Smelee, I. et al. NHG standaard astma bij volwassenen. Huisarts Wet. 50, 537–551 (2007).
4. Rabe, K. F., Vermeire, P. A., Soriano, J. B. & Maier, W. C. Clinical management of asthma in 1999: the asthma insights and reality in Europe (AIRE) study. Eur. Respir. J. 16, 802–807 (2000).
5. Chapman, K. R., Boulet, L. P., Rea, R. M. & Franssen, E. Suboptimal asthma control: prevalence, detection and consequences in general practice. Eur. Respir. J. 31, 320–325 (2008).
6. Gold, L. S., Smith, N., Allen-Ramey, F. C., Nathan, R. A. & Sullivan, S. D. Associations of patient outcomes with level of asthma control. Ann. Allergy Asthma Immunol. 109, 260–265 (2012).
7. van den Nieuwenhof, L. et al. Can the Asthma control questionnaire be used to differentiate between patients with controlled and uncontrolled asthma symptoms? A pilot study. *Fam. Pract.* **23**, 674–681 (2006).

8. Petsky, H. L., et al. Tailored interventions based on exhaled nitric oxide versus clinical symptoms for asthma in children and adults. *Cochrane Database Syst. Rev.* **16**, CD006340 (2009).

9. Petsky, H. L., et al. Tailored interventions based on sputum eosinophils versus clinical symptoms for asthma in children and adults. *Cochrane Database Syst Rev.* **18**, CD005603 (2007).

10. Borish, L. The immunology of asthma: asthma phenotypes and their implications for personalized treatment. *Ann. Allergy Asthma Immunol.* **117**, 108–114 (2016).

11. Halldar, P. et al. Cluster analysis and clinical asthma phenotypes. *Am. J. Respir. Crit. Care Med.* **178**, 218–224 (2008).

12. Hekking, P. P. & Bel, E. H. Developing and emerging clinical asthma phenotypes. *J Allergy Clin. Immunol. Pract.* **2**, 671–680 (2014).

13. Loza, M. J. et al. Validated and longitudinally stable asthma phenotypes based on cluster analysis of the ADEPT study. *Respir. Res.* **17**, 165 (2016).

14. Suijkerbuijk, A. W. M., et al. Maatschappelijke kosten voor astma, COPD en asthmo de allergie (Rijksinstituut voor Volksgezondheid en Milieu, 2012).

15. Green, L. A., Yawn, B. P., Lanier, D. & Dovey, S. M. The ecology of medical care revisited. *N. Engl. J. Med.* **344**, 2021–2025 (2001).

16. White, K. L., Williams, T. F. & Greenberg, B. G. The ecology of medical care. *N. Engl. J. Med.* **265**, 885–892 (1961).

17. OECD. *Health at a Glance* http://ec.europa.eu/health/reports/docs/health_glanc-e_2012_en.pdf (2012).

18. Juniper, E. F., O’Brien, P. M., Guyatt, G. H., Ferrie, P. J. & King, D. R. Development and validation of a questionnaire to measure asthma control. *Eur. Respir. J.* **14**, 902–907 (1999).

19. Juniper, E. F., Guyatt, G. H., Willan, A. & Griffith, L. E. Determining a minimal important change in a disease-specific quality of life questionnaire. *J. Clin. Epidemiol.* **47**, 81–87 (1994).

20. Newby, C. et al. Statistical cluster analysis of the British thoracic society severe refractory asthma registry: clinical outcomes and phenotype stability. *PLoS One* **9**, e102987 (2014).

21. Ortega, H., Miller, D. P. & Li, H. Characterization of asthma exacerbations in primary care using cluster analysis. *J. Asthma* **49**, 158–169 (2012).

22. Honkoop, P. J. et al. Asthma control cost-utility randomized trial evaluation (ACCURATE): the goals of asthma treatment. *BMC Pulm. Med.* **11**, 53 (2011).

23. RIVM. Gemiddelde body mass index (kg/m²) naar leeftijd en geslacht. (Rijksinstituut voor Volksgezondheid en Milieu, 2012).

24. World Health Organisation (WHO). BMI adults % overweight (2017). http://www.who.int/mediacentre/factsheets/fs311/en/.

25. National Health and Nutrition Examination Survey. in *Healthy Weight, Overweight, and Obesity Among U.S. Adults.* (ed CDC) (2017). https://www.cdc.gov/nchs/data/nhanes/gdataadultweight.pdf.

26. Honkoop, P. J. et al. Adaptation of a difficult-to-manage asthma programme for implementation in the Dutch context: a modified e-Delphi. *NPJ Prim. Care Respir. Med.* **27**, 16086 (2017).

27. Morrison, D. et al. Digital asthma self-management interventions: a systematic review. *J. Med. Internet Res.* **16**, e51 (2014).

28. Honkoop, P. J. et al. Symptom- and fraction of exhaled nitric oxide-driven strategies for asthma control: a cluster-randomized trial in primary care. *J. Allergy Clin. Immunol.* **135**, 682–688 (2015).

29. Heaney, L. G. & Robinson, D. S. Severe asthma treatment: need for characterising patients. *Lancet* **365**, 974–976 (2005).

30. Abdi, H. in *Encyclopedia for Research Methods for the Social Sciences* (eds Lewis-Beck, M., Bryman, A. & Futing, T.) 792–795 (Sage, Thousand Oaks, CA, 2003).

31. Ward, J. H. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **58**, 226–244 (1963).

32. NAEP. *Expert Panel Report 3: Guidelines for the Diagnosis and Management of Asthma* 328–362 (National Heart, Lung and Blood Institute, 2007). http://www.nhlbi.nih.gov.

33. Juniper, E. F., Bouquet, J., Abetz, L. & Bateman, E. D. Identifying ‘well-controlled’ and ‘not well-controlled’ asthma using the asthma control questionnaire. *Respir. Med.* **100**, 616–621 (2006).