Sex and BMI alter the benefits and risks of sulfonylureas and thiazolidinediones in type 2 diabetes: A framework for evaluating stratification using routine clinical and individual trial data

Running title: Stratification of therapy in type 2 diabetes

**Authors:** John M Dennis¹ (MSc), William E Henley¹ (PhD), Michael N Weedon² (PhD), Mike Lonergan³ (PhD), Lauren R Rodgers¹ (PhD), Angus G Jones⁴,⁵ (PhD), William T Hamilton² (MD), Naveed Sattar⁶ (FMedSci), Salim Janmohamed⁷ (PhD), Rury R Holman⁸,⁹ (FRCP), Ewan R Pearson³ (PhD), Beverley M Shields⁴* (PhD), Andrew T Hattersley⁴,⁵* (DM) on behalf of the MASTERMIND consortium

*Joint corresponding

**Affiliations:**

1. Health Statistics Group, University of Exeter Medical School, Exeter, U.K.
2. Institute of Biomedical & Clinical Science, University of Exeter Medical School, Exeter, U.K.
3. Division of Molecular & Clinical Medicine, Ninewells Hospital, Dundee, U.K.
5. Royal Devon and Exeter National Health Service Foundation Trust, Exeter, U.K.
6. Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, U.K.

**Corresponding authors**

1. Beverley M Shields, Institute of Biomedical & Clinical Science, RILD Building, Royal Devon & Exeter Hospital, Barrack Road, Exeter EX2 5DW, U.K. Email: {HYPERLINK "mailto:B.Shields@exeter.ac.uk" }; phone +44 1392 408203.
2. Andrew T Hattersley, Institute of Biomedical & Clinical Science, RILD Building, Royal Devon & Exeter Hospital, Barrack Road, Exeter EX2 5DW, U.K. Email: {HYPERLINK "mailto:A.T.Hattersley@exeter.ac.uk" }; phone +44 1392 408260.

**Word count:** abstract 292, manuscript 4228.

Tables: 2. Figures: 2
Abstract

OBJECTIVE

The choice of therapy for type 2 diabetes after metformin is guided by overall estimates of glycemic response and side-effects seen in large cohorts. A stratified approach to therapy would aim to improve on this by identifying subgroups of patients whose glycaemic response or risk of side-effects differ markedly. We assessed if simple clinical characteristics could identify patients with differing glycemic response and side-effects with sulfonylureas and thiazolidinediones.

RESEARCH DESIGN AND METHODS

We studied 22,379 patients starting sulfonylurea or thiazolidinedione therapy in U.K. Clinical Practice Research Datalink (CPRD) to identify features associated with increased one-year HbA1c fall with one therapy class and reduced with the second. We then assessed if pre-specified patient subgroups defined by the differential clinical factors showed differing five-year glycemic response and side-effects with sulfonylureas and thiazolidinediones using individual randomised trial data from ADOPT (first-line therapy, n=2,725) and RECORD (second-line therapy, n=2,222). Further replication was conducted using routine clinical data from the GoDARTS (n=1,977).

RESULTS

In CPRD male sex and lower BMI were associated with greater glycemic response with sulfonylureas and a lesser response with thiazolidinediones (both p<0.001). In ADOPT and RECORD non-obese males had a greater overall HbA1c reduction with sulfonylureas than thiazolidinediones (p<0.001); in contrast obese females had a
greater HbA1c reduction with thiazolidinediones than sulfonylureas (p<0.001).

Weight gain and oedema risk with thiazolidinediones were greatest in obese females however hypoglycaemia risk with sulfonylureas was similar across all subgroups.

**CONCLUSIONS**

Patient subgroups defined by sex and BMI have a different pattern of benefits and risks on thiazolidinedione and sulfonylurea therapy. Subgroup specific estimates can inform discussion about the choice of therapy after metformin for an individual patient. Our approach using routine and shared trial data provides a framework for future stratification research in type 2 diabetes.
In type 2 diabetes there is limited guidance to help clinicians and patients choose between the different glucose-lowering therapy options recommended after metformin. { ADDIN EN.CITE { ADDIN EN.CITE.DATA }} Guidelines suggest a discussion of the benefits, adverse effects, and costs of therapy to select the most appropriate medication for a particular patient. { ADDIN EN.CITE

Estimates of important clinical outcomes such as HbA1c, weight change and risk of side-effects are at present derived from whole trial populations and a key question is whether
they vary across patient subgroups defined by simple characteristics. { ADDIN EN.CITE


If estimates do vary by simple characteristics this may provide a starting point for a stratified approach in type 2 diabetes; the ‘targeting of treatments according to the biological or risk characteristics shared by patients’. { ADDIN EN.CITE { ADDIN EN.CITE.DATA }

Sulfonylureas and thiazolidinediones are recommended second and third line therapy options in all major type 2 diabetes guidelines. { ADDIN EN.CITE { ADDIN
They represented 50% of new second line prescriptions in 2016 in the U.S (sulfonylureas 46%, thiazolidinediones 4%). 

Glycemic response, weight change and common side effects have been well described in whole trial populations for both therapies. Differences in glycemic response by sex and BMI with thiazolidinediones and sulfonylureas have been previously suggested in observational studies.
systematically compared whether the benefits and risks of these therapies vary across subgroups defined by simple clinical patient characteristics.

Sulfonylureas and thiazolidinediones have, in contrast to newer therapies, been evaluated head-to-head in two long-term, randomized trials, ADOPT and RECORD.\{ ADDIN EN.CITE \{ ADDIN EN.CITE.DATA \}} ADOPT showed there was a greater durability of response up to 5 years with the thiazolidinedione rosiglitazone compared to either the sulfonylurea glyburide or metformin.\{ ADDIN EN.CITE \{ ADDIN EN.CITE.DATA \}} The full individual participant data of both trials are now available through Clinical Study Data Request,\{ ADDIN EN.CITE\}

and a current topic of debate is how to improve the output of secondary research projects using such shared trial datasets.\{ ADDIN EN.CITE\}
In this study we present a practical and cost-effective framework for stratification research using shared trial datasets alongside routine clinical data. We applied this framework to systematically evaluate whether simple clinical patient characteristics can be used to stratify therapy with sulfonylureas and thiazolidinediones.
Research Design and Methods

**Framework for stratification research**

In discovery analysis we explored routine clinical data to identify simple characteristics associated with glycemic response to sulfonylureas and thiazolidinediones, and used the results to define patient subgroups likely to show differential response. In validation analysis we evaluated differences in response within subgroups as a pre-specified hypothesis in ADOPT and RECORD, the two largest head-to-head randomized trials of sulfonylureas and thiazolidinediones available via Clinical Study Data Request. We also evaluated the secondary outcomes of weight change and risk of the common side effects of hypoglycemia, oedema and fracture within each subgroup (see Supplementary Figure 1 for our framework for stratification research using routine clinical and shared trial data).

**Datasets**

We analysed four datasets. Due to its large sample size, discovery analysis was conducted in routine clinical data from UK Clinical Practice Research Datalink (CPRD), with validation in trial datasets (ADOPT and RECORD) and a further routine clinical dataset (GoDARTs). Scientific approval for the use of CPRD data was granted by the CPRD Independent Scientific Advisory Committee (ISAC 13_177R) and permission to use the GoDARTs data was granted by the East of Scotland Regional Ethics Committee (09/21402/44). Data for both ADOPT and RECORD trials were accessed through the Clinical Trial Data Transparency Portal under approval from GSK (Proposal 930).
CPRD is the world's largest database of anonymized primary care electronic health records.\cite{Williams2012}

Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource\cite{Williams2012}.

The database allows researchers to conduct studies on various health topics using anonymized data from primary care practices in the UK. It has been widely used in numerous studies and collaborations across different fields.

\cite{Williams2012}
Our study protocol for CPRD data ascertainment has been previously reported. We studied 22,379 non-insulin treated patients with type 2 diabetes and prescription records for a sulfonylurea or thiazolidinedione from the February 2014 build of CPRD GOLD (see CPRD data supplement for product codes). We included patients with a duration of diabetes over one year (to minimise effect of lifestyle change following diagnosis).
and at least one year on-therapy without change in co-prescribed glucose lowering therapy (see Supplementary Figure 2 for CPRD patient flow diagram). { ADDIN EN.CITE

**Trials**

ADOPT and RECORD were prospective type 2 diabetes trials over at least 5 years of, respectively, glycemic durability and cardiovascular outcomes, in participants randomized to thiazolidinedione, sulfonylurea or metformin therapy. { ADDIN
In ADOPT we included participants in the intention to treat population with a valid baseline BMI randomized to sulfonylurea (glibenclamide) or thiazolidinedione (rosiglitazone) therapy (n=2,725). In RECORD we included participants in the intention to treat population on background metformin randomized to sulfonylurea (glibenclamide (18%), gliclazide (30%) or glimepiride (52%) according to local practice) or thiazolidinedione (rosiglitazone) add-on therapy (n=2,222).

**Genetics of Diabetes Audit and Research in Tayside Study (GoDARTs)**

GoDARTs contains information from the medical records of 18,276 people resident in eastern Scotland. We examined 1,977 patients with type 2 diabetes and valid prescription records for a sulfonylurea or thiazolidinedione.

**Analysis – data extraction and definitions**

**CPRD – discovery analysis**

The primary outcome was one year glycemic response in patients starting therapy with a sulfonylurea (any) or thiazolidinedione (pioglitazone or rosiglitazone) for the first time.

We extracted HbA1c at therapy start and at one year to calculate initial HbA1c response (one year HbA1c – baseline HbA1c; see CPRD data supplement for HbA1c codes), and baseline clinical characteristics: sex, BMI, age at diagnosis, duration of diabetes and eGFR.
Baseline HbA1c was defined as the closest HbA1c to the drug start date in the 91 days prior to the drug start date. One year HbA1c was defined as the closest HbA1c to one year after drug start date (+/-3 months). HbA1c response was only valid if there were no changes to diabetes medications between 60 days prior to the baseline HbA1c and the date of the one year HbA1c.
No adjustment was made for dose. To evaluate the secondary outcomes of long-term response and side effects we extracted measures of body weight, HbA1c, and records of fracture and oedema (see CPRD data supplement for fracture and oedema codes) over five years from the start of therapy. Patients with a fracture or oedema record in the two years prior to the drug start date were excluded from fracture and oedema analyses. We defined adherence as a Medication Possession Ratio (the number of days of available medication divided by the number of days between the first and last prescription dates, multiplied by 100). Due to the association between adherence and response, only patients issued sufficient prescriptions (medical possession ratio of between 80% and 120%) were included in analysis.

**Trials – validation analysis**
We used individual participant data from the trials to validate initial findings in CPRD. Based on the CPRD results we pre-specified four subgroups defined by sex and obesity (BMI > 30kg/m²). For each subgroup we compared average glycemic response by therapy over five years as the difference in area under the HbA1c response curve. This is equivalent to the time-updated HbA1c measure used in the UKPDS outcomes model. 

At years one, three and five we also estimated the difference between therapies in average glycemic response. We assessed annual weight change (percentage...
change from baseline) using the same approach. We also compared durability of response by therapy as measured by time to therapy failure. Failure was defined as in the original trials (ADOPT: confirmed fasting plasma glucose ≥180 mg/dl; RECORD confirmed HbA1c ≥8.5%). To evaluate side effects over five years we estimated the on-therapy risk of fracture (any), clinically determined peripheral oedema (all events, moderate/severe events (as defined as in the original trials as sufficient to, respectively, interfere with or prevent normal everyday activities)) and clinically determined hypoglycemia (all, moderate/severe as defined in the original trials). In ADOPT we excluded patients with a history of oedema from oedema analysis, in RECORD history of oedema was not available.

**Godarts**

We evaluated average glycemic response by therapy over five years using the same approach used for CPRD.

**Statistical Analysis**

**Short-term response: CPRD**

We assessed associations between baseline clinical characteristics (BMI, sex, age at diagnosis, duration of diabetes, eGFR) and one year glycemic response in linear regression models. A series of baseline HbA1c-adjusted models examined each clinical characteristic in turn, separately for each therapy.
Should Studies of Diabetes Treatment Stratification Correct for Baseline HbA1c?
We conducted a complete case analysis for each variable of interest, including all patients with valid data even if they had missing data for other clinical characteristics. Diagnostic plots of residuals were examined to check model assumptions were met. Based on the initial analysis we defined four subgroups defined by sex and obesity (BMI>30 v BMI<30 kg/m²) and for each therapy calculated baseline HbA1c adjusted least-square mean estimates of one-year response for each subgroup. To test for an overall effect of heterogeneity by sex and obesity subgroup we used a likelihood ratio test to compare a model with a drug:subgroup interaction with a nested model without an interaction term.

**Long-term response, weight gain and side effects: trial data**

We compared how each outcome was altered by therapy in each subgroup separately. We conducted response and weight change analysis in each trial separately, but pooled the data for side effects to increase study power. To estimate glycemic response over time we fitted baseline adjusted repeated measures mixed effect models using on-therapy HbA1c values at each study visit (n=22 ADOPT, n=19 RECORD) up to five years, including fixed effects for study visit, baseline HbA1c, therapy, visit by therapy interaction and visit by baseline HbA1c interaction, and patient-level random effects with an unstructured covariance matrix. Missing on-therapy HbA1c records were assumed to be missing at random. We calculated point estimates and 95% confidence intervals (CI) for the difference in average glycemic response by therapy at years one, three and five through contrasts of least-squares mean HbA1c change. We tested for an overall effect of heterogeneity by subgroup using the same interaction test as in CPRD. Weight change was modelled using the same approach.
To measure the net difference in HbA1c response between therapies we calculated the cumulative area under the HbA1c response curve (AUC) for each participant at every study visit using the trapezoidal rule. Participant AUC was then used as the outcome in repeated measures mixed effects models of the same structure as for glycemic response. A least-squares mean point estimate (95% CI) was calculated at year five to contrast overall response by therapy.

Time to therapy failure and side effects were estimated using the Kaplan-Meier method and Cox proportional hazards regression. Proportional hazards assumptions were evaluated using Schoenfeld residuals and were satisfied for all analyses. For each side effect the hazard ratio contrasting thiazolidinedione therapy with sulfonylurea therapy was estimated for each subgroup using an individual participant meta-analysis of data from both trials.

**Long-term response, weight gain and side effects: CPRD**

In CPRD we replicated analyses using the same models as described above for all outcomes except hypoglycemia, which is poorly captured in primary care records. For analysis of long-term HbA1c response, we extracted all HbA1c records between 60 days prior to the drug start date up to five years after the drug start date whilst on unchanged therapy. HbA1c records were categorised to three monthly intervals (nearest HbA1c record +/-1.5 months) to enable comparison with the trials. Where data points were missing, results were interpolated to ensure each time point reflected the same population of patients. The same approach was used for weight change, but with weights extracted at 6 monthly intervals (+/-3 months). For time to failure analysis, therapy failure was defined as two consecutive HbA1cs ≥8.5% or one HbA1c ≥8.5% followed by the addition of another therapy (the same definition of
glycemic failure used in RECORD). Data were censored if prescription records ended before a change in therapy. We excluded patients with changes to diabetes therapy without a prior HbA1c ≥8.5% as these changes were unlikely to relate to glycemic failure.

CPRD data extraction was conducted using Stata v13.0. All other analyses were conducted using R.
Results

**Routine clinical data: sex and obesity are associated with differential glycemic response with sulfonylureas and thiazolidinediones**

In CPRD we examined clinical factors associated with one year glycemic response amongst 22,379 eligible patients (10,960 thiazolidinedione; 11,419 sulfonylurea) (see Supplementary Table 1 for baseline characteristics). Sex and BMI showed the greatest differential response to therapy (Supplementary Figure 3). Compared to males, females had a greater response with thiazolidinediones, but a lesser response with sulfonylureas (both p<0.001). Higher BMI was associated with greater response with thiazolidinediones, but a lesser response with sulfonylureas (both p<0.001). Older age at diagnosis and lower eGFR were associated with a greater response to both therapies, there was greater response to thiazolidinediones with shorter diabetes duration, and greater response to sulfonylureas with longer diabetes duration and higher HDL (Supplementary Figure 3).

As sex and BMI showed the greatest differential response we specified four subgroups defined by sex and obesity (BMI>30 v BMI<30kg/m²) for use in subsequent analysis. We found evidence of heterogeneity of response by subgroup (p<0.001). Figure 1 shows one year glycemic response by therapy for the four subgroups. Non-obese males had a greater one year response with sulfonylureas than thiazolidinediones (baseline adjusted change in HbA1c: -13.2 v -9.7 mmol/mol, p<0.001), whereas obese females had a greater one year response with thiazolidinediones than sulfonylureas (-13.8 v -9.4 mmol/mol, p<0.001). Obese males and non-obese females showed similar responses with both therapies (both p=0.6).
Results were consistent for pioglitazone and rosiglitazone when analysed separately, and for gliclazide and non-gliclazide sulfonylureas (Supplementary Figure 4).

**Trial data: non-obese males have greater glycemic response with sulfonylureas, obese females with thiazolidinediones**

We went on to assess if the sex and obesity defined subgroups also showed differential response when randomly allocated to therapy in the ADOPT (n=2,725) and RECORD (n=2,222) trials. Randomisation resulted in well matched patients for each therapy within each subgroup (see Supplementary Tables 2 and 3 for baseline characteristics). There were marked differences in response with both therapies in the four subgroups with a clear similarity between the two trials (test for heterogeneity in ADOPT and RECORD both p<0.001, Figure 2a and 2b). Over five years there was a greater overall glycemic response for non-obese males with sulfonylureas (both trials p<0.001), relating to the greater earlier benefit with sulfonylureas over thiazolidinediones that persisted beyond 2 years in both trials. In contrast there was a greater overall glycemic response for obese females with thiazolidinediones over sulfonylureas (both trials p<0.001), and there was little early benefit with sulfonylureas.

**Trial data: absolute risk of therapy failure differs markedly by subgroup**

We assessed the risk of monotherapy and dual-therapy failure, respectively, in ADOPT and RECORD. In both trials for non-obese males there was no difference in the five year risk of failure on the two therapies but all other subgroups were less likely to fail with thiazolidinediones than sulfonylureas (Hazard ratios 0.23-0.72, test for heterogeneity ADOPT p<0.001, RECORD p=0.01, Table 1, Supplementary Figures 5-6). In ADOPT, risk of monotherapy failure at five years with thiazolidinediones was lower for obese females (11%) than non-obese males (22%),
whilst with sulfonylureas failure risk was lower for non-obese males (22%) than obese females (42%) (Table 1).

**Trial data: increased risk of weight gain and oedema with thiazolidinediones for all subgroups**

Weight was increased for all subgroups with thiazolidinediones compared to sulfonylureas but this was much more marked in obese females (Figure 2c, Supplementary Figure 7). Oedema was more common with thiazolidinediones compared to sulfonylureas for all subgroups; this resulted in the largest difference in absolute risk for obese females who are most likely to develop oedema regardless of therapy (Table 2, Supplementary Figure 8).

**Trial data: increased risk of fracture with thiazolidinediones only for females**

Fracture was more common with thiazolidinediones compared to sulfonylureas but only for females. Absolute risk was similar for obese and non-obese females (Table 2, Supplementary Figure 9).

**Trial data: increased risk of hypoglycemia with sulfonylureas for all subgroups**

Sulfonylureas, compared with thiazolidinediones, increased the risk of moderate/severe hypoglycemia for all subgroups (Table 2, Supplementary Figure 10). Hazard ratios for hypoglycemia of any severity were consistent with those for moderate/severe events (Supplementary Tables 4-5). For all side effects there was a similar differences between therapies when the trials were analysed separately (Supplementary Tables 4-5).
Routine clinical data: results for long-term glycemic response, time to failure and side effects were consistent with trial data

In CPRD and GoDARTs (see Supplementary Table 6 for GoDARTs baseline characteristics), five year glycemic response results were consistent with the trials (Supplementary Figures 11 & 16). In CPRD differences by therapy in time to failure results were similar to the trials although absolute failure rates were higher (Supplementary Figure 12). Weight gain, oedema and fracture results in CPRD were comparable to trial data (Supplementary Figures 13-15).

Summary of results

For subgroup data summaries of glycemic response, weight change and risk of side effects estimates specific to each sex and obesity defined subgroup see Supplementary material.
Conclusions

Stratification of therapy with sulfonylureas and thiazolidinediones is possible using sex and BMI

We have robustly demonstrated across four datasets that sex and BMI alter the benefits and risks of type 2 diabetes therapy with sulfonylureas and thiazolidinediones. We show in non-obese males the glycemic response with sulfonylureas is better on average in the first 5 years than on thiazolidinediones, without excess weight gain, but with an increased risk of hypoglycemia. For obese females there is a clear glycemic benefit over the first 5 years with thiazolidinediones compared to sulfonylureas, but there is increased weight gain and susceptibility to oedema and fracture. Our findings will allow for much more informed discussion of the benefits and risks of these therapies than the present ‘one size fits all’ approach (see supplementary Subgroup Data Summary for estimates specific to each sex and obesity defined subgroup).

Our results provide the first example of stratification of therapy in type 2 diabetes based on simple patient characteristics.
A recent data-driven cluster analysis proposed five subgroups of diabetes with differing disease progression and risk of complications, but did not evaluate whether subgroups differed in their response to therapy.
To date successful stratification in other conditions has involved expensive genetic testing, as applied in cancer and single gene diseases such as monogenic diabetes.
Expensive testing is unlikely to become practical or justified in type 2 diabetes, a highly prevalent condition with relatively inexpensive therapy. Type 2 diabetes genetic studies have identified polymorphisms associated with drug response but the impact of these, at present, are too small to guide clinical management, in contrast to our results.

A framework for stratification research using shared trial data alongside routine clinical data

This study is an early and important demonstration of how shared trial data can be harnessed to meaningfully benefit patients.
We propose a novel and cost-effective framework to use shared trial data in stratification research. Our framework can be applied to study other type 2 diabetes therapies and to study stratification in other chronic conditions. It has great potential to improve the output of future studies using shared trial data.

**Comparison to previous studies**

Whilst no existing studies have systematically assessed whether both the benefits and risks of these two therapies are altered by clinical characteristics, previous analyses have suggested sex and BMI are associated with glycemic response to both therapies. In ADOPT, risk of therapy failure were lower for obese and female subgroups with thiazolidinediones compared to sulfonylureas, but an interaction was not tested for and the difference in glycemic trajectory was not examined. Increased response for obese female patients with thiazolidinediones and for male patients with sulfonylureas has been found in observational studies but the impact of this in terms of stratification has not been assessed. We have previously shown that markers of insulin resistance including BMI are associated with reduced glycemic response to DPP4 inhibitors but not glucagon-like peptide 1 (GLP-1) receptor agonists but evidence for other agents is limited.

Previous studies have also found sex and BMI alter the risk of side effects. The increase in fracture risk with thiazolidinediones applies mainly to post-menopausal women and is consistent within trials. We found hypoglycemia risk with sulfonylureas was similar across subgroups even
though glycemic response differed, and this needs further investigation. Whilst our study shows absolute risk of oedema with thiazolidinediones was highest in the obese female subgroup that had the greatest response, further study is required to fully evaluate the association between glycemic response and the risk of common side effects for these therapies.

**Limitations**

Our study has limitations. The results do not allow prediction at an individual level, however we present subgroup estimates that will better reflect the likely outcome for an individual patient within that subgroup than outcome estimates derived from whole trial populations. Rosiglitazone, the thiazolidinedione used in both trials analysed in our study, has been withdrawn in many countries due to concerns over cardiovascular safety.

Routine clinical data support a thiazolidinedione class effect of differential response by sex and obesity but trial data were not made
available to repeat our analysis for pioglitazone. Previous meta-analyses suggest that the risks of oedema and fracture are similar with both drugs, further supporting the generalizability of our findings to pioglitazone. For sulfonylureas, a similar pattern of results was observed in ADOPT (glibenclamide), RECORD (52% glimepiride, 30% gliclazide, 18% glibenclamide), and routine clinical data (including a gliclazide only analysis), supporting a sulfonylurea class effect. In CPRD, for the one year glycemic response analysis we excluded non-adherent patients and those whose anti-hyperglycemic therapy was altered (potentially due to poor response, very good response, or poor tolerance) within the first year, and this could have accounted for the differences we observed when comparing sulfonylurea and thiazolidinedione therapy. However, we saw a similar pattern of glycemic response differences using time to failure and mixed effect models which both included all patients with at least one on-therapy HbA1c measure for up to five years. The CPRD time to failure analysis was also limited as patients whose treatment was intensified below the HbA1c failure threshold of 8.5% were censored rather than defined as experiencing therapy failure. A strength of the CPRD analysis is the demonstration of consistent results with all three analytical approaches, each with their own strengths and weaknesses. Measured or unmeasured baseline differences between patients could have explained findings in the routine data, but are very unlikely to explain the differences we observed in the randomized clinical trials, further highlighting the strength of our study design. Over 90% of patients in the datasets studied were White Caucasian, limiting the applicability of our findings to other racial groups, a common problem with trials in type 2 diabetes. Additional data would be required to answer whether there are differences in patients of South Asian, Hispanic or Black origin, where fat
distribution can be markedly different and a different obesity cut-off may be appropriate. { ADDIN EN.CITE 
Cite> <Year>2004</Year> <RecNum>111</RecNum> <DisplayText>(40) 
The ideal stratified approach would be based on cardiovascular endpoints rather than the intermediary measure of glycemic response. In this analysis we were underpowered to detect differences for cardiovascular outcomes in RECORD (the primary trial analysis showed no difference between rosiglitazone and sulfonylureas or metformin), or rarer side effects such as heart failure. Given the two recent trials demonstrating cardiovascular benefits with SGLT2 inhibitors and GLP-1 receptors agonists each required over 7000 high-risk participants, it may be that impractically large trials are required for stratification of cardiovascular endpoints.

**Future research**

Evaluation of the risks and benefits of newer therapies such as DPP4 inhibitors, SGLT2 inhibitors and GLP-1 receptor agonists will require routine clinical data from large numbers of patients alongside shared head-to-head drug trial data, and will be possible in the near future. Given the greater expense of newer therapies cost-effectiveness evaluation will be necessary in this work. The ongoing GRADE study will give important long-term head-to-head comparative effectiveness data on second-line treatment with insulin, DPP4 inhibitor, GLP-1 receptor agonist and sulfonylurea therapy.
Further mechanistic studies are required to interrogate the mechanisms underlying differential response to sulfonylureas and thiazolidinediones. Thiazolidinediones act through the adipocyte and so it is likely any increase in the number of adipocytes will improve glycemic response. This provides a potential explanation for our findings as women, compared to men, have more adipocytes as they have a higher whole body
percentage fat mass and obese subjects have more adipocytes than non-obese subjects. { ADDIN EN.CITE

The reduced insulin sensitivity seen in obesity is likely to explain the reduced response to sulfonylureas that predominantly stimulate insulin secretion by the beta-cell. The consistently better response seen in males to sulfonylureas was unexpected and further studies are required to define the mechanism of this observation.

Clinical Implications

The sex and obesity subgroup-specific estimates presented in this study will allow a much more informed discussion between clinicians and patients of the benefits and risks of sulfonylureas and thiazolidinediones, at no cost. We recommend this discussion with an individual patient is based around the appropriate sex and obesity subgroup-specific estimates presented for the two therapies in the Subgroup data summary (Supplementary material). Whether this alters a decision on therapy will depend on the individual circumstances of the patient, as the trade-off between early-response, long-term durability and risk of side-effects will be different.

Conclusion
Simple patient characteristics alter the benefits and risks of type 2 diabetes therapy with sulfonylureas and thiazolidinediones. Our novel and practical framework for stratification research can be applied in type 2 diabetes and other chronic conditions, and has great potential to improve output from future studies using shared trial data.
Funding

The MASTERMIND consortium is supported by the Medical Research Council (UK) (MR/N00633X/1). ATH and RHH are NIHR Senior Investigators. ERP is Wellcome Trust New Investigator (102820/Z/13/Z), ATH is a Wellcome Trust Senior Investigator. ATH and BMS are supported by the NIHR Exeter Clinical Research Facility. AGJ is supported by an NIHR Clinician Scientist award. NS acknowledges support by Innovative Medicines Initiative Joint Undertaking under grant agreement no. 115372, the resources of which comprise financial contribution from the European Union’s Seventh Framework Programme (FP7/2007-2013) and EFPIA companies’ in kind contribution. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The funders had no role in any part of the study or in any decision about publication.

Conflict of Interest statement

WEH declares a grant from Quintiles, ERP declares personal fees from Lily, Novo Nordisk, and Astra Zeneca. NS declares personal fees from Boehringer Ingelheim, Eli Lilly, Novo Nordisk, Janssen and a grant from Astra Zeneca. RHH declares research funding from Bayer, Astra Zeneca, MSD and honoraria from Amgen, Bayer, Elcelyx, Janssen, Intarcia, MSD, Novartis, Novo Nordisk and Servier. SJ is an employee and stockholder of GSK. Representatives from GSK, Takeda, Janssen, Quintiles, AstraZeneca and Sanofi attend meetings as part of the industry group involved with the MASTERMIND consortium. No industry representatives were involved in the writing of the manuscript or analysis of data. For all authors these are outside the submitted work; no other relationships or activities that could appear to have influenced the submitted work.

Author Contributions

ATH, ERP, BMS, JMD, AGJ, SJ, WTH and WEH designed the study. BMS, LRR and MNW extracted the data from the CPRD. JMD, BMS and ML analysed the data. JMD and BMS drafted the article. ATH, ERP, NAS, WEH, AGJ and RHH provided support for the analysis and interpretation of results, and critically revised the article. All authors approved the final article. ATH and BMS are the guarantors of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation

Parts of this study were presented in abstract form at the Diabetes UK Professional Conference, U.K., March 2016, the American Diabetes Association’s Scientific Sessions, New Orleans, June 2016, and the Annual Meeting of the European Association for the Study of Diabetes, Munich, September 2016.
References

{ ADDIN EN.REFLIST }
Table 1: Absolute and relative risk of glycemic failure with thiazolidinediones (TZD) and sulfonylureas (SU) in trial data, by sex and obesity defined subgroup. Failure defined according to original trial protocol (ADOPT trial (monotherapy) defined as fasting plasma glucose $\geq$ 180mg/dl; RECORD (dual therapy with metformin) defined as HbA1c $\geq$ 8.5%. Relative risks presented as hazard ratios (95% CI) for TZD compared to SU.

<table>
<thead>
<tr>
<th>ADOPT monotherapy failure</th>
<th>No. of patients</th>
<th>No. of events</th>
<th>Absolute 5 year risk (%)</th>
<th>Hazard ratio (95% CI) (TZD vs. SU)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TZD</td>
<td>SU</td>
<td>TZD</td>
<td>SU</td>
<td>TZD</td>
</tr>
<tr>
<td>Non-obese males</td>
<td>373</td>
<td>395</td>
<td>47</td>
<td>63</td>
<td>21.7%</td>
</tr>
<tr>
<td>Obese males</td>
<td>402</td>
<td>387</td>
<td>44</td>
<td>108</td>
<td>15.0%</td>
</tr>
<tr>
<td>Non-obese females</td>
<td>208</td>
<td>174</td>
<td>16</td>
<td>34</td>
<td>10.9%</td>
</tr>
<tr>
<td>Obese females</td>
<td>407</td>
<td>379</td>
<td>31</td>
<td>93</td>
<td>11.6%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>RECORD dual-therapy failure</th>
<th>No. of patients</th>
<th>No. of events</th>
<th>Absolute 5 year risk (%)</th>
<th>Hazard ratio (95% CI) (TZD vs. SU)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TZD</td>
<td>SU</td>
<td>TZD</td>
<td>SU</td>
<td>TZD</td>
</tr>
<tr>
<td>Non-obese males</td>
<td>240</td>
<td>228</td>
<td>66</td>
<td>70</td>
<td>33.6%</td>
</tr>
<tr>
<td>Obese males</td>
<td>361</td>
<td>356</td>
<td>92</td>
<td>132</td>
<td>30.7%</td>
</tr>
<tr>
<td>Non-obese females</td>
<td>137</td>
<td>127</td>
<td>26</td>
<td>45</td>
<td>20.7%</td>
</tr>
<tr>
<td>Obese females</td>
<td>379</td>
<td>394</td>
<td>72</td>
<td>142</td>
<td>22.7%</td>
</tr>
</tbody>
</table>
Table 2: Absolute and relative risk of side effects over 5 years with thiazolidinediones (TZD) and sulfonylureas (SU) in ADOPT & RECORD combined, by sex and obesity defined subgroup. Relative risks presented as hazard ratios (95% CI) for TZD compared to SU. HRs and p values from meta-analysis of both trials.

<table>
<thead>
<tr>
<th>Side effect</th>
<th>No. of patients</th>
<th>No. of events</th>
<th>Absolute 5 year risk (%)</th>
<th>Hazard ratio (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TZD</td>
<td>SU</td>
<td>TZD</td>
<td>SU</td>
<td>TZD</td>
</tr>
<tr>
<td>Non-obese males</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oedema (Moderate/Severe)</td>
<td>607</td>
<td>620</td>
<td>13</td>
<td>4</td>
<td>3%</td>
</tr>
<tr>
<td>Fracture (All)</td>
<td>613</td>
<td>623</td>
<td>26</td>
<td>18</td>
<td>7%</td>
</tr>
<tr>
<td>Hypoglycaemia (Moderate/Severe)</td>
<td>613</td>
<td>623</td>
<td>14</td>
<td>90</td>
<td>3%</td>
</tr>
<tr>
<td>Obese males</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oedema (Moderate/Severe)</td>
<td>740</td>
<td>719</td>
<td>37</td>
<td>16</td>
<td>7%</td>
</tr>
<tr>
<td>Fracture (All)</td>
<td>763</td>
<td>743</td>
<td>30</td>
<td>28</td>
<td>6%</td>
</tr>
<tr>
<td>Hypoglycaemia (Moderate/Severe)</td>
<td>763</td>
<td>743</td>
<td>13</td>
<td>70</td>
<td>2%</td>
</tr>
<tr>
<td>Non-obese females</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oedema (Moderate/Severe)</td>
<td>340</td>
<td>293</td>
<td>13</td>
<td>5</td>
<td>5%</td>
</tr>
<tr>
<td>Fracture (All)</td>
<td>345</td>
<td>301</td>
<td>31</td>
<td>8</td>
<td>14%</td>
</tr>
<tr>
<td>Hypoglycaemia (Moderate/Severe)</td>
<td>345</td>
<td>301</td>
<td>10</td>
<td>44</td>
<td>4%</td>
</tr>
<tr>
<td>Obese females</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oedema (Moderate/Severe)</td>
<td>749</td>
<td>746</td>
<td>60</td>
<td>25</td>
<td>10%</td>
</tr>
<tr>
<td>Fracture (All)</td>
<td>786</td>
<td>773</td>
<td>77</td>
<td>33</td>
<td>14%</td>
</tr>
<tr>
<td>Hypoglycaemia (Moderate/Severe)</td>
<td>786</td>
<td>773</td>
<td>18</td>
<td>83</td>
<td>3%</td>
</tr>
</tbody>
</table>
Figure legends

Figure 1: CPRD: One year glycemic response (baseline adjusted change in HbA1c) with thiazolidinediones (red dots) and sulfonylureas (blue triangles), by sex and obesity defined subgroup. Data are presented as least square means adjusted for baseline HbA1c ± 95% CI. A reduction (improvement) in HbA1c is represented as a negative value.
Figure 2: 5 year glycemic response (change from baseline in HbA1c) and weight change (percentage change from baseline) with thiazolidinediones (TZD, red dots) and sulfonylureas (SU, blue triangles), by sex and obesity defined subgroup. Data are presented as means at each study visit ± standard error from mixed effects models. A reduction (improvement) in HbA1c is represented as a negative value. For AUC and treatment difference estimates positive values favour SU, negative values favour TZD. For RECORD weight change data see Supplementary Figure 7.
a) ADOPT trial: absolute glycemic response (mmol/mol)

- **Non-obese males**
  - 5-year AUC difference T2D vs SU: -7.1 (99% CI: -2.7 to 11.5) mmol/mol
  - Treatment difference T2D vs SU: mmol/mol (99% CI)
  - 1 year: +3.9 (1.6 to 6.3); p=0.01
  - 3 years: +16 (10.4 to 21.8); p=0.01
  - 5 years: +33 (1.8 to 24); p=0.79

- **Obese males**
  - 5-year AUC difference T2D vs SU: -3.09% CI: -0.1 to 1.9) mmol/mol
  - Treatment difference T2D vs SU: mmol/mol (99% CI)
  - 1 year: +1.8 (1.1 to 3.1); p=0.01
  - 3 years: -1.6 (0.5 to 2.6); p=0.11
  - 5 years: -2.0 (-6.0 to 0.9); p=0.045

- **Non-obese females**
  - 5-year AUC difference T2D vs SU: -8.1 (99% CI: -11.1 to 6.5) mmol/mol
  - Treatment difference T2D vs SU: mmol/mol (99% CI)
  - 1 year: -3.5 (+0.3 to 6.3); p=0.01
  - 3 years: +14 (-0.4 to 15); p=0.001
  - 5 years: +23 (-3.9 to 3.8); p=0.001

- **Obese females**
  - 5-year AUC difference T2D vs SU: -18.6 (99% CI: -23.3 to 15.3) mmol/mol
  - Treatment difference T2D vs SU: mmol/mol (99% CI)
  - 1 year: +5.0 (+0.7 to 9.3); p=0.01
  - 3 years: +5.1 (-3.9 to -3.9); p=0.001
  - 5 years: -7.2 (-5.0 to -5.7); p=0.001

---

**RECORD trial: absolute glycemic response (mmol/mol)**

- **Non-obese males**
  - Weight change difference T2D vs SU (% 95% CI)
  - 1 year: -1.0% (-2.0% to -0.1%); p=0.04
  - 3 years: -0.5% (-1.9% to 0.9%); p=0.36
  - 5 years: +1.4% (0.3% to 2.6%); p=0.02

- **Obese males**
  - Weight change difference T2D vs SU (% 95% CI)
  - 1 year: -1.1% (-2.9% to 0.7%); p=0.08
  - 3 years: -0.5% (-1.9% to 1.0%); p=0.68
  - 5 years: +2.3% (0.9% to 3.7%); p=0.001

- **Non-obese females**
  - Weight change difference T2D vs SU (% 95% CI)
  - 1 year: +1.8% (0.8% to 2.5%); p=0.03
  - 3 years: +0.4% (0.0% to 3.5%); p=0.01
  - 5 years: -2.5% (0.0% to 4.9%); p=0.06

- **Obese females**
  - Weight change difference T2D vs SU (% 95% CI)
  - 1 year: +1.4% (0.0% to 2.7%); p=0.24
  - 3 years: +0.5% (0.3% to 3.7%); p=0.001
  - 5 years: +7.1% (5.5% to 8.8%); p<0.001

---

{ PAGE \* MERGEFORMAT }
b) ADOPT trial: weight change from baseline (%)